U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 240

Can OI Replace Hough? (Part I)

John D. Stackpole
Development Division

JULY 1981

Abstract

A preimplementation evaluation, in the form of an intercomparison between Hough analyses and forecasts and OI analyses and forecasts leads to the conclusion that replacing the operational Hough analysis with the OI at this time would not be advisable.

I. Introduction

During the late winter and spring of 1981, the Systems Evaluation Branch of the Development Division, NMC, conducted a series of preimplementation tests to address the question Can the Optimum Interpolation (OI) analysis system replace the Hough function system as the analysis system for the NMC Operational (or Large Scale) forecast model run? (The operational model is the Sela-spectral model with global 30 wave resolution and 12 vertical layers (SMG3C)). Early results, obtained in a developmental context by the Atmospheric Analysis Branch, (Dey and Morone, personnel communication) indicated that such a replacement would be feasible and beneficial to NMC's operations.

Following established procedures a preimplementation test was designed to take a close look at the question in the larger context of the major operational uses of the analysis. What follows is a description of the test design, a fair sampling of the results, and a discussion leading to the conclusion that replacing the Hough analysis at this time would not be advisable.

II. Test Design

Side by side comparison of analysis systems (rather than data reduction methods or forecast models) presents some unique problems. The simple-minded approach of asking which system fits the data better is not a fair or valid test - data have errors of measurement; an analysis system should recognize this and not fit all the data closely. On the other hand, of course, an analysis independent of the data would not be particularly desirable either. Thus it becomes important that experienced meteorologists make a subjective evaluation of the analysis, particularly in any context where the analysis, per se, is important to NMC's operations.

The principal use of any analysis is, of course, the establishment of initial conditions for a forecast. Thus the effect of the differing analyses upon the same forecast model becomes the main evaluation test, and the majority of the evaluations center on the relative merits of the two forecasts.

In brief outline the contending analysis systems are:

A. The Hough Analysis

This is a spectral analysis method in which a set of Hough functions are fitted to the observations in a least squares sense, along with empirical orthogonal functions of pressure. This system has been in long use at NMC; the fundamentals are described by Flattery (1971).

B. The Optimum Inrerpolation Analysis

This is an analysis to gridpoints system which has the particular advantage of taking account of the known or presumed accuracies of observing systems. It also has the advantage of taking place within the sigma-coordinate system of the forecast model, thus avoiding an interpolation from pressure coordinates. The most recent description of the current system is by McPherson (1980).

In outline the test is quite straight forward: the OI analysis is run using the same (operational) data base as was previously (in the regular operations) used by the Hough analysis. The analyses are then used as input to the same forecast model and various results are compared.

Both analyses use the same first guess - the twelve hour forecast from the FINAL Data Assimilation run.

The Hough analysis, since it is in pressure coordinates, requires an interpolation into sigma coordinates, the OI does not. The interpolation is a portion of the initialization of the spectral model which of course is bypassed for the OI analysis initial conditions.

Once the initial sigma data were obtained, everything was identical, almost. Both forecasts started out with the 12 layer 30 wave version (SMG3C) with a 2 iteration, 4 mode nonlinear normal mode initialization, and a forecast to 48 hours. At 48 hours the horizontal resolution was reduced to 24 waves and the forecast proceeded to 84 hours. At that point, the vertical resolution was reduced from 12 to 6 layers (the horizontal resolution remained at 24 waves) and the forecast proceeded to 264 hours (11 days). Nothing, however, is ever perfect. After the tests were completed we discovered that the forecasts based on the OI analyses ("OI forecasts" for short) were run with mountains that had but 24 waves of resolution while the Hough forecasts had 30 wave mountains. Thus the OI forecasts might have been under a little disadvantage up to 84 hours. The differences between the 24 and 30 wave mountains are slight, though, and it seems very unlikely that this difference could have an effect of any significance.

(This all came about because the first quess from the 24 wave FINAL is the source of the mountains for the OI forecasts but not for the Hough forecasts).

The evaluation of the analyses centered naturally on their operational uses.

The analyses themselves serve, of course, as the pictorial representation of the current state of the atmosphere; they had better be "reasonably" close to the observational data and also look "right" (meteorologically sound) or else we are in all sorts of trouble.

At present the analyses also serve as persistence forecasts (for aviation) in the Southern Hemisphere and tropics to 20°N or so. Pending is a proposal to start issuing the Southern Hemisphere forecasts instead of the analyses. The Southern Hemisphere evaluation got somewhat compounded as a result: Should the OI replace the Hough analysis persistence forecasts and/or should either the Hough based or OI based forecasts replace the persistence all together? (There is good evidence, available elsewhere, that the Hough forecasts are indeed better than the persistence analyses in the tropics and Southern Hemisphere).

Since the main use of the analyses is for the forecasts, the major portion of the evaluation centers on the principal uses of the forecasts. There are four categories of these:

1. Marine Forecasts

The concern here is for the areas of the Atlantic and Pacific not covered by the LFM forecasts. The concentration is on the synoptic features at the surface and 500 mb for 24 and 48 hours. The subjective evaluator (D. Saxton of Forecast Division's Basic Weather Branch) was given these maps:

Sea Level pressure and Thickness	24 & 48 hr.
500 mb Height and Vorticity	24 & 48 hr.

2. Quantitative Precipitation

We live in hope that eventually the forecast model will get good at forecasting precipitation amounts - D. Olson (Forecast Division, QPB) looked at

SLP & Thickness	24 & 48 hr.
500 mb Ht. & Vorticity	24 & 48 hr.
Mean RH & 700 mb V.V.	24 & 48 hr.
QPF	24 & 48 hr.

to see if the OI analysis made any difference in the forecasts over the U.S.

3. Aviation Weather

One of the major and most critical users of the forecasts is the aviation weather community, both for flight planning and inflight weather problems. Here also the analyses are important both for short range "forecasting" and because they serve as the persistence forecasts in the Southern Hemisphere. R. McCarter (Aviation Weather Branch, Forecast Division) was given, for subjective evaluations;

Northern and Southern Hemisphere

Polar stereographic maps of:

o SLP & Thickness	Analysis & 24 hr fcst
o 500 mb ht & Vorticity	Analysis & 24 hr fcst
o 250 mb ht & isotachs	Analysis & 24 hr fcst
o 100 mb ht & isotherms	Analysis & 24 hr fcst

Tropical Mercator Strip maps of:

o 250 mb winds and stream function	Analysis & 24 hr fcst

4. Medium Range Forecast

The longer range forecasts (past 48 hours) from the model serve as one of the primary guidance tools for the Medium Range Forecast group in their efforts to divine the future in the 3 to 5 day (and 6 to 10 day) ranges. F. Hughes was given maps of

- o SLP & Thickness — Days 3, 4, and 5
- o 500 mb ht & ht change — Days 3, 4, and 5
- o 5 day mean 500 ht and departure from normal — Centered on day 8

and he undertook both the subjective and objective evaluations of the differences.

A fill-the-blanks questionnaire accompanied each of the sets of maps to help the evaluators focus their attention and to aid in summarizing their opinions. Appendix A shows the form that was used.

Objective verifications, where the analyses and forecasts were compared to observations, were also performed. Details will be found in a later section.

III. Execution of the Tests

The 5 cases deemed adequate on the basis of Dey's earlier results were picked on an ad hoc basis by the attendees at NMC's regular map discussion. The cases were generally selected because something of interest was going on and there was the possibility, at least, that a change in the analysis might make a difference for good or ill. The case dates were:

#1 12Z 30 Jan 81 (Mid continent snow storm)

#2 00Z 10 Feb 81 (good forecast of big storm)

#3 12Z 4 Mar 81

#4 00Z 17 Mar 81

#5 00Z 18 Apr 81

#4 and 5 were selected on the basis of medium range forecast interest. A sixth case, 00Z 6 May 81, was selected for medium range forecast verification only - the MRF people pay attention to 00Z initial time cases only, and the initial set of 00Z cases was a little sparce.

Other than the usual problems of working with large computer systems and complex programs no particular difficulties were encountered. There were some unanticipated delays in obtaining Southern Hemisphere graphics products and, as usual, the tightest bottleneck was the Varian processing.

One piece of "inadvertent experimentation" (i.e. an error caught and corrected in re-runs) was performed: Three of the cases were run without benefit of upper air bogus in the OI analysis. Although no detailed study was made it was apparent in the with-bogus reruns that the presence or absence of the bogus data had little or no impact on the forecasts. The current rational for this somewhat distressing observation is that the normal mode initialization is in effect rejecting the single level heights-only bogus data. What to do about this remains a problem for another time.

IV. Objective Evaluations

A. Height, Temperature and Wind

It is not possible to tailor a set of objectively calculated verification statistics to any particular set of relevant subjective evaluation criteria; all we can do is attempt to look at a sufficiently wide assortment of statistics and scores under the assumption that severe difficulties or substantial differences will manifest themselves. Indeed past experience (Stackpole et. al., 1978; Stackpole, 1980) has shown this to be the case; subjective and objective judgements of differences (where there are any of significance) closely parallel one another.

The objective statistics computed comprised the now familiar set: Tewles/Wobus S1 Score (Brown, 1971), the mean (bias) and rms errors of geopotential heights, temperatures, winds and relative humidity computed at the mandatory levels 850 mb, 500 mb, 250 mb, and 100 mb.

The analyses (following normal mode initialization in the forecast model's sigma coordinates and interpolation back to pressure), and forecasts at 24, 48, and 84 hours were verified against sets of observations in both the northern and southern hemisphere. Unfortunately not all of the cases could be verified for all of the time periods because of missing or exceedingly sparse data, particularly in the southern hemisphere.

In the northern hemisphere four observation networks were used: NH102 (102 stations quasi-uniformly distributed over the hemisphere), NA110 (110 stations over North America), and two networks in the tropical (0°- 30°N) north Atlantic and Pacific, SNATL48 and SNPAC45. In the southern hemisphere three networks were used: SH31 (31 quasi-uniform stations), AUS24 (24 stations regularly reporting, mainly at 0000 GMT, from Australia), and SAM30 (30 stations in South America, all south of the equator, and reporting mainly at 1200 GMT).

Excepting the S1 score, the method of calculation of the various error statistics is straightforward. For the mean and rms error statistics, the analysed and forecast quantities were biquadratically interpolated to the station locations, the errors established and the appropriate summations, over all the stations with valid reports in the network, were performed.

For the S1 score calculation, a preliminary pass was made through all of the available upper air observations (not just those of the network in question) and the station which was the nearest neighbor to each of the network stations was located. Then the observed and forecast height gradients between the station pairs are used for the S1 score calculation. The "nearest neighbor" selection is limited by claustrophobic (pairs closer than 100 km are not allowed), agoraphobic (pairs separated by more than 2357 km are not allowed) and geminiphobic (if A selects B as its closest neighbor, B may not select A) constraints. This method of calculation of S1 differs from the usual one in which the gradients are computed between pre-selected grid points in a fixed geographic array, and results in S1 numbers that are slightly larger as a consequence.

For the objective verification of precipitation forecasts a different network of 60 first order stations (long in use by NMC Forecast Division) was augmented by 30 additional stations designed to fill some gaps and cover problem areas in coastal and mountain areas. A computer algorithm was readied, designed to interpolate (in a manner appropriate to the discontinuous precipitation fields) from the grid points at which precipitation was forecast to these stations, thus producing a list of 12 hour accumulated precipitation amounts for the verification times and stations. These station forecasts were the material for the calculation of the precipitation threat and bias scores.

The medium range forecasts were objectively verified (by the forecasters concerned) with statistics appropriate to the different nature of the forecasts - pattern correlations for the departure from normal maps and Heidke Skill scores (relative to climatology) for the accumulated precipitation.

A representative selection of height, temperature and wind verifications are seen in Figure 1 through 18, in the form of scatter or "impact" charts. In these charts the ordinate is the error of the analysis or forecast from the Optimum Interpolation method, and the abscissa the error from the Hough function method. Thus, points (each representing one particular case) falling below and to the right of the 45° diagonal (for the positive-definite statistics: S1 and rms) signify "OI better" and vice versa. For the bias errors the "domains of betterness" are slightly more involved but obvious from the figures. "Betterness" is defined as the state of having the magnitude of the error closer to zero. On the figures the verification networks are distinguished by the figure title, the particular statistic and level is indicated in the body of the figure and the time range (analysis through 84 hour forecast) by the plotted character. The units are metric; S1 is dimensionless.

The analysis "errors" are included in the figures as a sort of safety check - one does not expect analysis programs to draw to the data exactly, particularly, as is the case here, when the "analysis" is the output from a normal mode initialization process. However any large discrepancy in the two analysis systems would give one pause. Such does not seem to be the case here, thankfully.

Turning to the figures then, we see in Figure 1 that as far as the mid and upper tropospheric heights (over the northern hemisphere) go, it matters little which analysis is used to initiate the forecast. All of the points cluster about the no impact line with something of an increase in dispersion (and error level) the further out in time one goes. The temperature forecast errors, Figure 2, tells a similiar story, with some variations. The tropospheric temperatures show an "it-makes-no-difference" progression both in the mean and rms error plots while the 100 mb diagram shows the OI analysis to be a better fit to the data (in the rms sence) but with a slight positive bias. This rms better fit is maintained throughout the forecasts while the tendency for the model to cool the stratosphere brings the mean error points into the region of "OI better", too, as time goes on. This latter has nothing particular to do with the analyses, a corresponding mean warming, without any effect on the impact, can be seen in the 850 mb mean error plot. The 250 mean temperatures are remarkably well behaved in the model.

Figure 3, the 250 and 100 mb wind error plots, brings us back to the no-impact category again. The 250 mb speeds are a mite slow in the Hough analysis (and a mite fast at 100 mb) and the forecasts show no major changes or pronounced patterns. The rms vector errors, show nothing more than a no-impact decrease in forecast quality no matter which initial analysis is used.

Figures 4 through 6 follow the same pattern for the NA110 network and lead to essentially the same conclusions. The temperature errors, both mean and rms, show somewhat more case-by-case variability than did the hemispheric ones, and the OI's advantage in terms of 100 mb rms errors is not so apparent, but the overall results remain the same. The same statements hold for the wind errors, too.

Figures 7 through 12, for the set of northern hemisphere tropical stations in the Atlantic (mainly in the Carribean and south to the equator along the South American coast) and the Pacific, are something of a dissapointment. On a priori grounds one could expect a better representation of the atmosphere from the O/I because of various constraints built into the Hough function system. If such an improvement is present, it is not manifesting itself in any dramatic way in the forecasts. The height errors scatter neatly about the no impact line as, for the most part, do the wind errors. There is perhaps some suggestion of positive impact from the OI in the 100 mb wind speed errors - the scatter is rather large, however, and the magnitude of the errors suggests that the "impact" is one of changing a lousy forecast into merely a poor one. That's something anyway. Better representation can't overcome the defects of inadequate observations in the first place. There is one clear victory for the O/I, however, in the 100 mb temperature biases (the rms errors are controlled by the large biases and have little independent information). The forecasts from either analysis show a warm bias (this is related to the lack of resolution in the model sigma coordinates in the vicinity of the high tropical tropopause) but the O/I bias is substantially and consistently less than the Hough. This is presumably accountable for by recalling that the O/I analysis takes place in the model sigma layers and hence the information suffers from but one interpolation in returning to constant pressure surfaces, while the Hough pressure level analysis information must be interpolated twice, once into the model and then (after the forecast) out again. The effect of these double interpolations shows up, it seems, only in the tropical tropopause region. Elsewhere the resolution is adequate to its task.

When we turn to the Southern hemisphere, Figures 13 through 18, we see things are somewhat more chaotic: the figures show more scatter and slightly larger error values on balance than in the north. However no clear pattern, other than "no difference", emerges from the figures. There is a tendency for the 100 mb temperatures to show a preference for the OI analysis based forecasts, in the Australian/South American networks - this is most likely because those networks reach well into the tropics and would be influenced by the same tropopause resolution difficulties as in the northern hemisphere tropics.

The (very) sharp-eyed reader may have noticed that many of the southern hemisphere diagrams do not have a full complement of points - five cases were run, thus there could be a maximum of five circles, dots, etc. Points were not plotted for two reasons: either 1) too few reports were available (for whatever reason) for the particular synoptic time to generate meaningful statistics; or 2) the point(s) fell outside the limits of the figures as sketched. In the latter case the scales could have been changed to accomodate the extreme points, with an accompanying loss of resolution, but there seems little point in displaying distinctions between forecasts that are either horrible or just awful. Also in most cases the extreme points fell rather close to the no-impact line. Both forecasts were in the lousy class.

The general conclusion from the objective height, temperature, and wind verifications is that there is no, or at best minimal, significant difference between the Hough and O/I analyses insofar as they effect the quality of the forecasts out to 84 hours. Where there are differences they favor the OI analysis slightly.

A potentially serious problem with the OI analysis is, implicitly, visible in the objective verification statistics. One can see, particularly in the 100 mb temperature and wind verifications, that a set of points shows a considerably poorer skill than the other cases, at each of the forecast times. (In some cases the point fell outside the plotted chart). Generally the positioning of the points places them in the "Hough Better" region of the diagrams. Although the charts don't show this, the extreme points all came from the same forecast, the initial time of 00Z 18 Apr 81. (Two of the subjective verifiers most strongly brought this case to my attention as well). The cause of the difficulties lies in the first guess for both the OI and Hough analyses. The first guess for these, as for all the cases, is the OI/Spectral forecast from the Global data assimilation cycle (GDAC) (the "FINAL") operationally run at NMC. Table 1 shows the number of days preceeding the forecast case dates that the GDAC ran uninterruptedly:

Date	Days
12Z 30 Jan 81	7
00Z 10 Feb 81	4
12Z 4 Apr 81	2
00Z 17 Mar 81	3
00Z 18 Apr 81	10

Table 1. Number of Uninterrupted Cycle Days Preceeding Test Date

(An interruption means that the operational Hough analysis replaces the GDAC forecast as the first guess). It is obvious that the offending case had been preceeded by a considerably longer span of OI cycles than the others. There was a substantial build up of noise, a higher levels, in the OI cycles by this time and it passed on through the OI analysis of 00Z 18 Apr 81 to make problems in the forecasts. The Hough analysis, with its inherently greater smoothing, reduced but didn't entirely eliminate the problem. This problem had been recognized by the analysis people independently; a heavy diffusion was added to the stratospheric sections of the model used in the GDAC on 20 April. However if this or some other device is not adequate to control the noise generation in an uninterrupted GDAC series, replacing the Hough analysis with the OI could present appreciable problems for the system as a whole. It has been, after all, the intermittent interjections of the Hough analyses that have kept the GDAC noise under control.

Objective Evaluations

B. Precipitation

Figure 19 is a (compact) summary of the precipitation verifications of the five cases: on the left side are figures relating to the areal (rain/no rain) verification (every 12 hours) for the western and eastern portions of the country divided by the eastern limit of the Rocky Mountains, and on the right, quantitative verification scores. The precipitation threat scores, Tsp, (the intersection of the sets of stations with observed and forecast rain occurance, divided by their union) show a reasonably parallel behavior in time of the forecasts from the two analyses with the exception of the 36 hour OI based forecast (more on this below). It appears that the OI forecasts pull ahead of (in the west) or catch up with (in the east) the Hough forecasts in the latter portions of the forecasts. This is coupled with a decrease in the bias (as implied in the lower left figures) of the OI forecasts relative to the Hough - a better threat score and lower bias is an auspicious combination. The quantitative threat scores and indications of rainfall amount (either totaled over the five cases or averaged per station) make for a somewhat less coherent pattern. The threat shows the Hough based forecasts as clearly better in the east - in the west the quantitative biases are right on the button and the threats are mixed. (The subjective evaluation of the forecasts - see below - suggests that not too much should be made of these differences - they represent small differences between not very good forecasts).

However, because of the importance of precipitation we show here a few examples that will illustrate some of the HUF/OI differences and also the general difficulty the model has (in its present configuration) in forecasting precipitation no matter what initial conditions it starts with.

Figure 20 shows two twelve hour forecasts (of 12 hour accumulated amounts) and the verifying observations for one of the cases. The Hough based forecast looks o.k. as far as areal coverage goes (although the model fails to catch the heavy, presumably convective, rain of the south-eastern states) but the O/I forecast is considerably dried out. This is not atypical - the O/I humidity analysis is capable of paying closer attention to the horizontal and vertical gradients of moisture than is the Hough; presuming this to be a "better" analysis, in that it follows the data closer, the problem may lie in some inability of the forecast model to respond properly in the short range. Figure 21 illustrates the same problem in a 36 hour forecast in a west coast context. The Hough forecast captures only a portion of the California - Oregon rain but that's better than the OI did. Finally Figure 22 puts the problem in an interior context. The OI/Hough differences in Montana and Saskatchewan favor the Hough forecast while both forecasts fared poorly over the south central area.

In the light of these particular case results the general summary of Figure 19 presents something of a puzzle: the improvement of the OI forecasts relative to the Hough at the longer ranges, particularly in the west. About all that can be said is to point out the phenomenon and remark that as the influence of the less well analysed Pacific region ("less well" for lack of data) comes in over the verification areas, some subtle advantage of the OI analysis comes to the fore. It would seem well worth looking further into this in an effort to improve the shorter range forecasts.

Subjective Evaluations

The subjective evaluator's opinions of the relative merits (or demerits) of the forecasts are most conveniently presented as numerical tabulations - one point each for whichever analysis based forecast was prefered (or tie). The tables are arranged in the same format as the questionnaire in the appendix.

1. Basic Weather (D. Saxton)

	24 hr.			48 hr.			Total		
SLP & Thickness	OI	HUF	TIE	OI	HUF	TIE	OI	HUF	TIE
East Pacific & Alaska --------	1	2	2	2	2	1	3	4	3
Western (West of 105°) U. S.--	1	1	3	1	1	3	2	2	6
Eastern U. S. ----------------	1	0	4	1	1	3	2	1	7
Western Atlantic -------------	0	0	5	1	2	2	1	2	7
500 mb Height & Vorticity									
East Pacific & Alaska ---------	0	0	5	1	3	1	1	3	6
Western U. S. ----------------	2	0	3	0	3	2	2	3	5
Eastern U. S. ----------------	0	0	5	0	0	5	0	0	10
Western Atlantic -------------	0	1	4	1	1	3	1	2	7
Totals ------	5	4	31	7	13	20	12	17	51

It is clear from these evaluations that D. Saxton found the differing analyses methods to produce forecasts of nearly comparable value with a slight imbalance favoring the Hough. Not that the forecasts were in all cases similar; indeed in notes on some of the cases Saxton remarked that one analysis produced a superior forecast in one geographic area while another area (in the <u>same</u> forecast) favored the other forecast/analysis. No clear pattern was to be seen, either set of forecasts would serve as suitable guidance. Not so for the precipitation forecasts ---

2. Precipitation (D. Olson)

	24 hr.			48 hr.			Total		
Rain/No Rain Coverage	OI	HUF	TIE	OI	HUF	TIE	OI	HUF	TIE
Western (West of 105°) U. S.--	0	3	2	1	2	2	1	5	4
Eastern U. S. ----------------	0	1	4	2	1	2	2	2	6
Quantitative Amounts									
Western U. S. ----------------	1	1	3	2	0	3	3	1	6
Eastern U. S. ----------------	0	0	5	1	1	3	1	1	8
Relative Humidity Patterns									
Western U. S. ----------------	1	1	3	1	1	3	2	2	6
Eastern U. S. ----------------	1	2	2	2	1	2	3	3	4
Utility of Mass/Motion Forecasts to QPF									
All U. S. --------------------	1	0	4	2	2	1	3	2	5
Totals -----	4	8	23	11	8	16	15	16	39

Although this tabulation gives the impression of tied results overall, (and this is correct in a sence), the remarks that accompanied the returned questionnaire put this result in a somewhat different light.

Words such as "both versions missed", "so bad", "wretched as usual", and "BAD!!" gave me the impression that D. Olson was not entirely satisfied with the quality of the precipitation forecasts from the model. In short, he said, they stunk. Not very good precipitation forecasts from the Spectral model were remarked upon in the earlier tests (Stackpole, 1980); clearly the model is not yet capable of responding meaningfully to whatever differences may exist in the analysis of moisture. The results are tied, certainly, but not for <u>first</u> place.

Olson had particularly strong comments about the 00Z 18 Apr case, pointing out vertical velocity centers in the OI based forecasts that were unrelated to relative humidity or precipitation in the forecast. These centers were also unrelated to centers in the Hough based forecasts. Presumably this is related to the noisy first guess that both analyses started with and also suggests that the effects of the noise were not confined to the upper atmosphere exclusively. The Hough analysis included an inherent filtering effect, the OI had considerably less filtering, if any.

3. Aviation (R. McCarter)

Considerable difficulty was experienced by Mr. McCarter in performing what he considered to be an adequate evaluation of the analysis and forecasts of the upper air parameters. The difficulties centered on two separable problems - one was the unanticipated (by me) unavailability of observational data in a format suitable for easy comparison with the analyses and forecasts. This was remedied by preparation of the 250 mb and 100 mb charts with both the contoured fields and plots of the "verifying" data on them. Mr. McCarter was then asked to pass judgement on these. He did so and the results are tabulated below. The other problem was that of the tropopause and tropopause wind shear forecasts - here the forecasts from either analysis had a sufficient number of defects that using them to select between analyses seemed pointless. The problem, as with the QPF, seems to lie in the model, not the initial conditions. The tropopause is constructed from the mandatory level forecast temperatures (and the wind shear from the winds) - necessary arbitrary choices need further refinement (such is in process) to generate more satisfactory tropopause fields.

This tabulation of the evaluations:

	ANAL			24 HR FCST			TOTAL		
	OI	HUF	TIE	OI	HUF	TIE	OI	HUF	TIE
250 mb hts and Isotachs (and Jet)									
U.S.	3	1	1	1	2	2	4	3	3
Pacific.	3	1	1	2	2	1	5	3	2
South America. . . .	2	1	2	1	1	3	3	2	5
Australia.	2	1	2	2	1	2	4	2	4
100 mb Hts & Isotherms									
Pacific.	1	0	4	1	3	1	2	3	5
U.S.	0	0	5	1	3	1	1	3	6
Atlantic	0	1	4	1	3	1	1	4	5
South America. . . .	1	2	2	0	2	3	1	4	5
Australia.	2	2	1	1	1	3	3	3	4
Totals	14	9	22	10	18	17	24	27	39

shows an overall (slight) preference for the OI analysis which is effectively balanced by a preference for the Hough based forecasts. A slightly closer look at the tabulation shows that the preference for the OI analysis was centered at the 250 mb pressure while the preference for the Hough forecasts arose from consideration of the 100 mb fields.

The preference for the OI analysis may have been induced by the test design in part - recall that the "analyses" evaluated were output from the sigma-coordinates at the initial time. Thus the Hough analyses, originally in pressure coordinates, had undergone a pressure-to-sigma-to-pressure interpolation. While the OI had only under gone a sigma-to-pressure change. The Hough thus experienced at least some smoothing which may have degraded its appearance in the eyes of the evaluator.

The highly problematic case of 00Z 18 Apr is included in the tabulation (and contributed to the HUF prefered tallies). It also, deservedly, was the subject of some rather pointed remarks by the evaluator drawing attention to the deficiencies of the noisy first guess and the effect thereof upon the analysis and forecast.

On the related matter of the use of forecasts to replace persistence in the Southern Hemisphere and tropics, the evaluator came to the conclusion that either forecast did a better job than the analysis for the winds, but he had strong reservations about the value of the temperature forecasts. (Independent statistical studies have lead to the same conclusions).

4. Medium Range Forecasts (Hughes)

No questionnaire was utilized, instead we show a tabulation of the statistical verifications computed by F. Hughes and some of his subjective remarks, case by case.

<u>Medium Range</u>

Case	84 hr	108 hr	132 hr	5 Day Mean D+8
<u>00Z 10 Feb</u> Subjective Preference	HUF	HUF	OI	HUF (but both fcsts "pretty poor")
<u>00Z 17 Mar</u> SLP Correlation	HUF = 40 OI = 50	19 38	-4 2	
500 mb ht Standardized Correlation				HUF = -16 OI = 4

OI better but differences hard to see subjectively; D+8 poor forecast.

Case	84 hr	108 hr	132 hr	5 Day Mean D+8
<u>00Z 18 Apr</u>				
SLP Correlation	HUF = 64 OI = 75	58 67	42 41	
500 mb ht Standardized Correlation				HUF = 48 OI = 26
<u>00Z 6 May</u>				
SLP Corelation	HUF = 59 OI = 57	32 38	32 14	
500 mb ht Standardized Correlation				HUF = 42 OI = -2

D+8 5 Day Precip Skill Score

HUF = 19.0
OI = 14.7

HUF clearly better particularly at important longer ranges.

The general conclusion from these results that the Hough forecasts are on balance, better than the OI is hardly a statistical certainty, based as it is on four cases. It also goes counter to the (subjective) conclusions reached by Dey and Morone, with only three cases to look at. The safest conclusion is that no firm statement can be fairly made; one is still left with a feeling of discomfort, however, over the prospect of the OI replacing the Hough at this point, as far as the impact on the medium range program is concerned.

V. Operational Considerations

The NMC operational sequence is quite time and deadline critical and at present has very little free time available. Thus any substantial increase in running time (wall clock time) for the analysis program could present serious problems. At present the Hough analysis requires about 15 minutes to complete its task in a protected operational framework. To this must be added about one more minute to accomplish the pressure to sigma conversions: A grand total of 16 minutes. Here is a table of the CPU and checkout wall times for the six test cases:

Case #	CPU	Wall
1	18'48"	31'37"
2	16'12"	54'44"
3	14'30"	36'10"
4	15'24"	56'49"
5	15'56"	27'14"
6	16'57"	38'14"

The wall times appear alarming but are misleading - the people concerned with the OI in the GDAC report that the CPU to wall time ratio is nearly one in the protected environment. Thus there seems little to be concerned about as far as OI execution times are concerned.

Furthermore a revised version of the OI analysis code runs some 4 minutes faster than the current one used in these tests. This code is undergoing the final stages of testing and checking prior to its introduction into the GDAC in NMC's FINAL cycle.

VI. Summary, Problems, & Conclusions

The overall results of this preimplementation test are something of a mixed bag: In the U.S. area and for the short range the general conclusion is that no really major difference can be seen. Saxton indicated a slight preference for the Hough based forecasts. There is some suggestion that the 18 Apr case was the principal contributor to this preference. That case was anomalous and could be discounted except that the conditions that led up to the anomaly are themselves important.

In the tropics there was also (dissapointingly) little difference. The OI did show up to good advantage in the higher atmosphere, presumably because of fewer vertical interpolations.

Similarly in the Southern Hemisphere, there was little upon which to base a choice of one analysis method over another.

At the 3 to 5 to 10 day range the rather limited sample precludes any firm conclusions but the weight of evidence tends to favor the Hough analysis based forecasts.

The 18 Apr "anomaly" remains a point of concern - the ideal is, of course, an uninterrupted run of the GDAC; until the noise generation problems have been put to rest one would be loath to eliminate the option of inserting occasional Hough analyses to quiet things down.

Other problems with the OI system which cropped up in the FINAL cycle while the current tests were underway (and not impinging directly on the tests) are also points of concern. An attempt to change the resolution from 24 to 30 waves in the GDAC ran into difficulties and had to be withdrawn; the faster running analysis code also ran into difficulties; there are continuing investigations of normal mode initialization and its impact on single level observations and the mass motion balance. The intent here is not to catalogue the trials and tribulations of the analysis developers (indeed solutions are in hand for most of the recognized problems) but to suggest that the OI system needs a little more seasoning in the operational framework before extending its domain further.

If we absolutely had to replace the Hough with the OI analysis, my judgement is that we could do so without any undue harm to NMC's operations, but there would be some uncomfortable moments and crises from time to time. Since we have the luxury of time (for now anyway) the recommendation is to remain with the Hough analysis. When the current goals of the OI in the FINAL are reached: fast code, 30 wave resolution, no excess noise, many uninterrupted trouble free cycles, well behaved normal mode initialization, etc., then it will be time to take another (presumably less extensive) look at the prospects for replacement.

REFERENCES

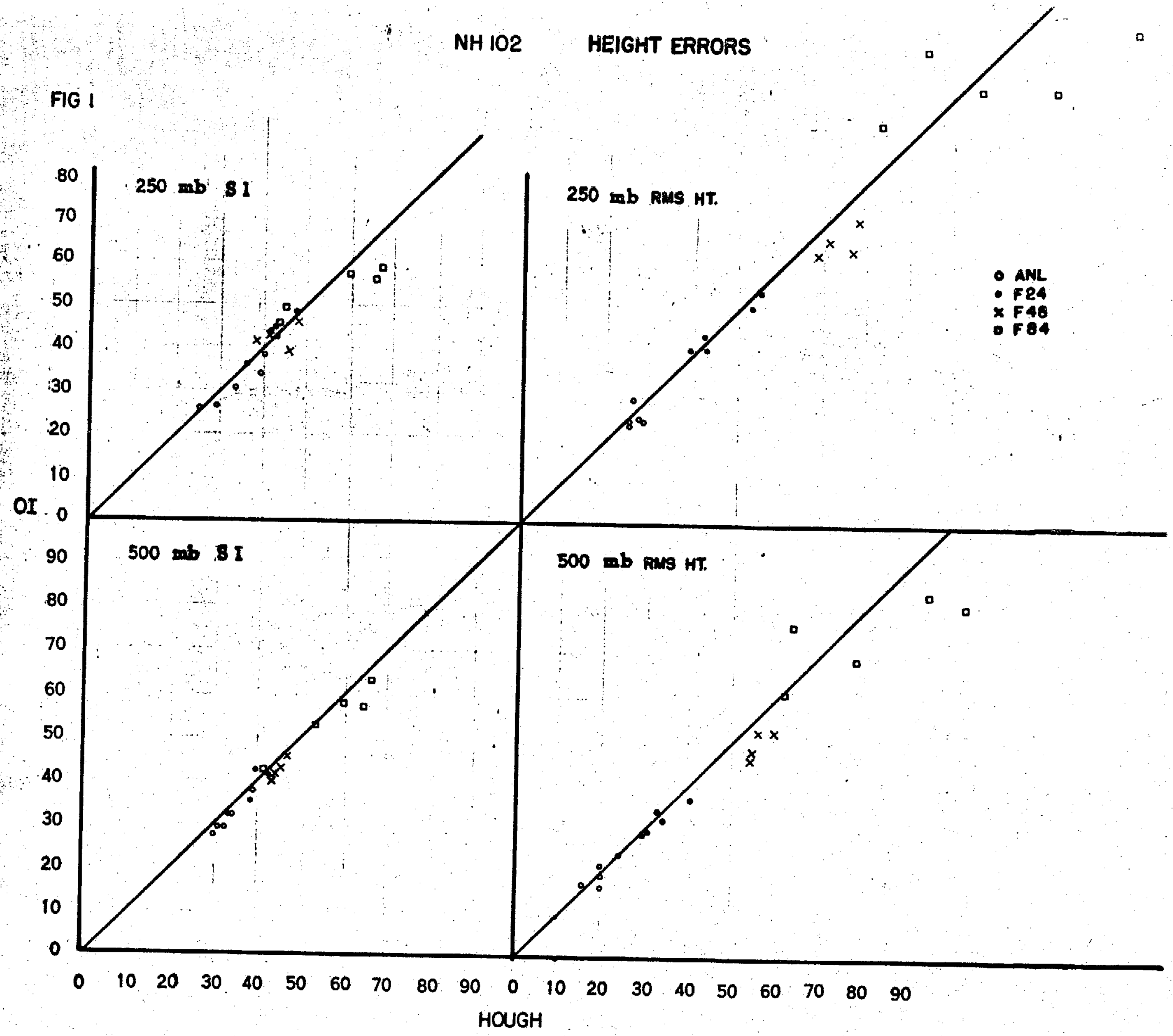
Brown, H., 1971: "Tewles-Wobus S1 Score". NMC Technical Attachment #71-2. 1 June 71. NMC, NWS, NOAA.

Stackpole, J. D., et. al., 1978: "How to Pick a New Forecast Model". Preprint Volume, Conference on Weather Prediction and Analysis and Aviation Meteorology, AMS, Boston MA, October 1978.

Stackpole, J. D., 1980: "How to Pick Another New Forecast Model (Spectral vs 7L PE)". NMC Office Note #215, NMC, NWS, NOAA.

Flattery, T. W., 1971: "Spectral Models for Global Analysis and Forecasting." Air Weather Service Technical Report #243, pp. 42-54.

McPherson, R. D., 1980: "NMC Office Notes 216, 217, 218, 219, 220, and 221. NMC, NWS, NOAA.



NH 102
HEIGHT ERRORS
FIG 1
250 mb S1
250 mb RMS HT.
500 mb S1
500 mb RMS HT.
o ANL
• F24
x F48
□ F84
OI
HOUGH
0 10 20 30 40 50 60 70 80 90

FIG 2

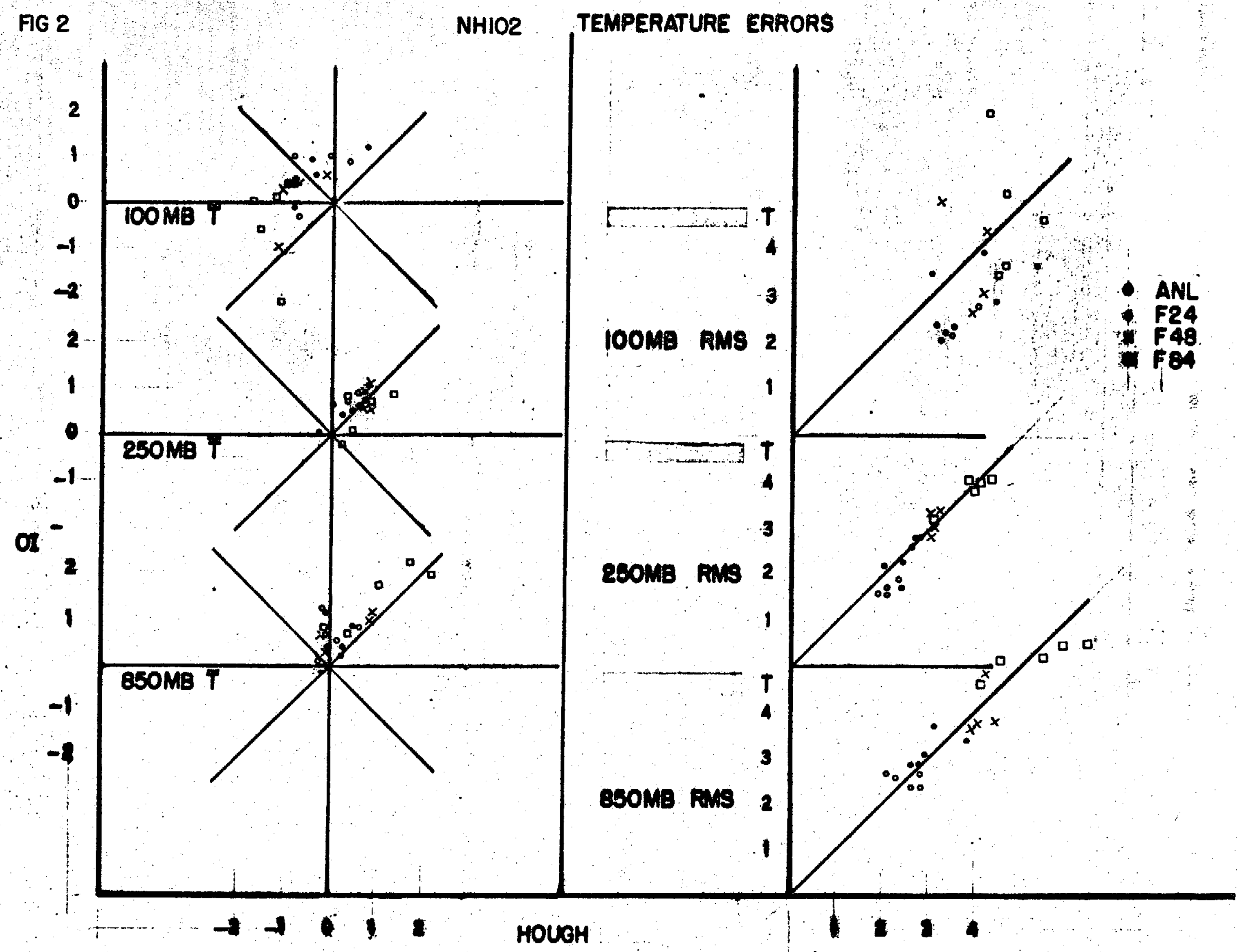

NH102 TEMPERATURE ERRORS
100MB T
250MB T
850MB T
100MB RMS
250MB RMS
850MB RMS
OI
HOUGH
ANL
F24
F48
F84

FIG 3

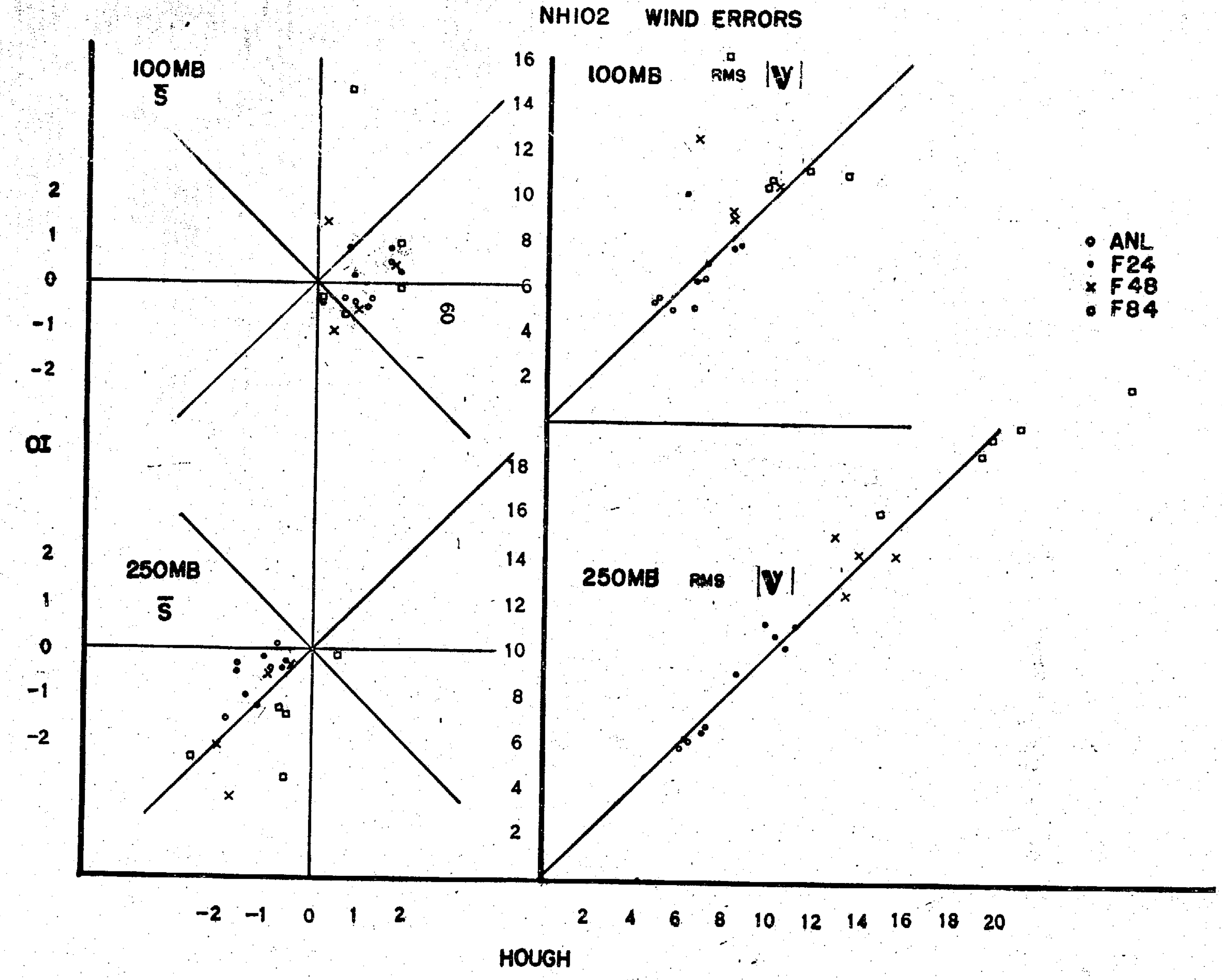

NH102 WIND ERRORS
100MB
$\bar{S}$
100MB RMS |V|
250MB
$\bar{S}$
250MB RMS |V|
OI
HOUGH
ANL
F24
F48
F84

FIG 4

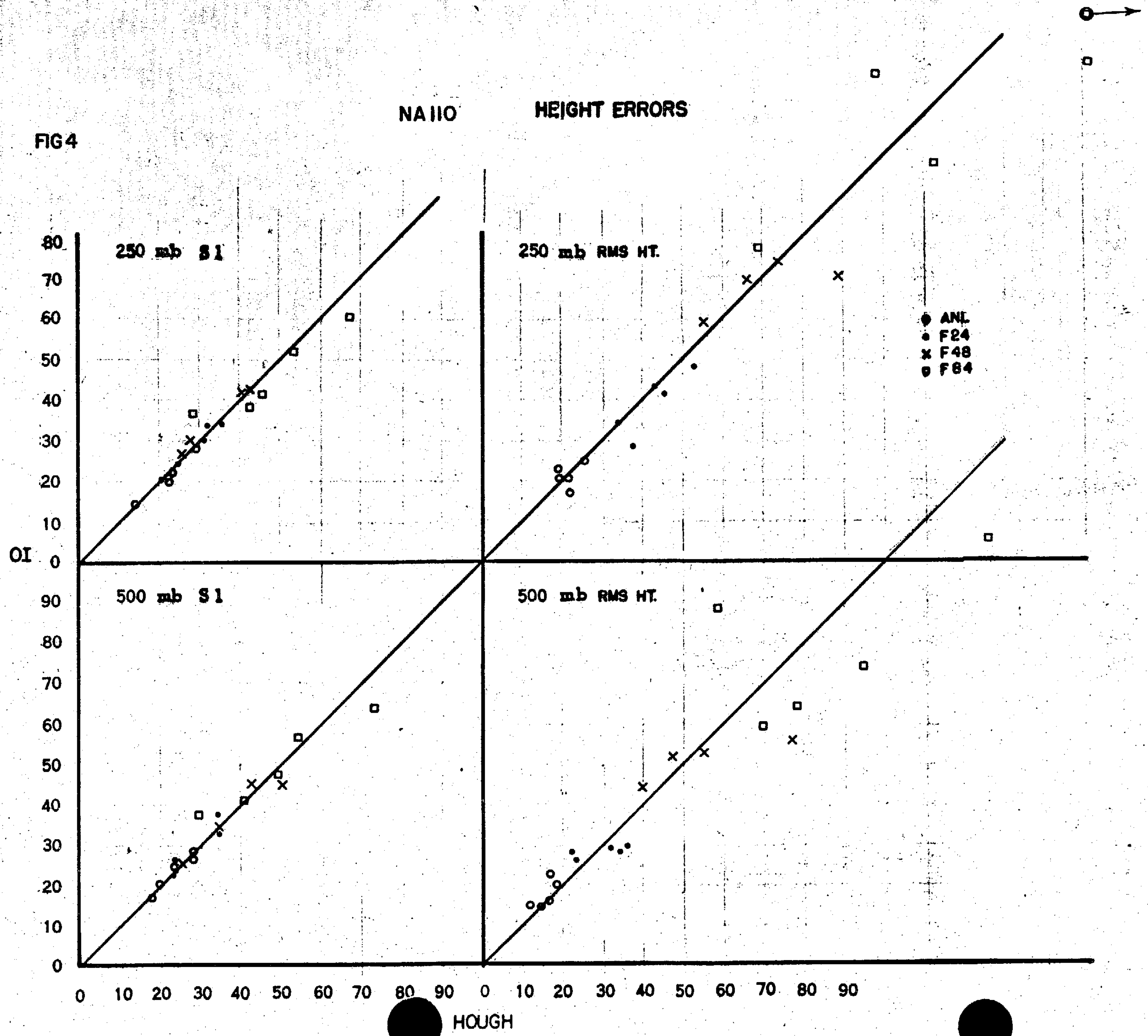

NA 110
HEIGHT ERRORS
250 mb S 1
250 mb RMS HT.
500 mb S 1
500 mb RMS HT.
ANL
F24
F48
F84
OI
HOUGH
0
10
20
30
40
50
60
70
80
90

FIG 5

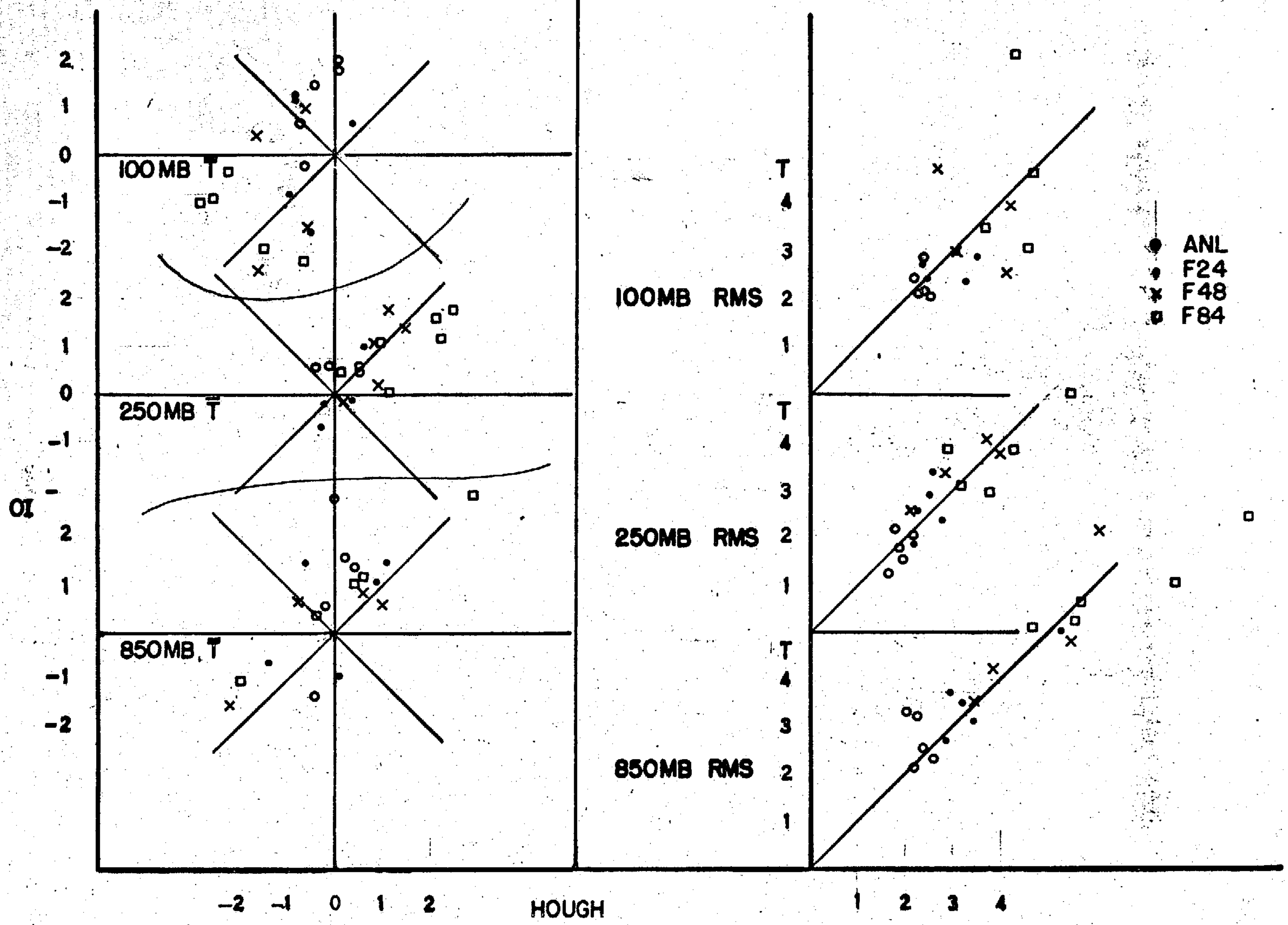
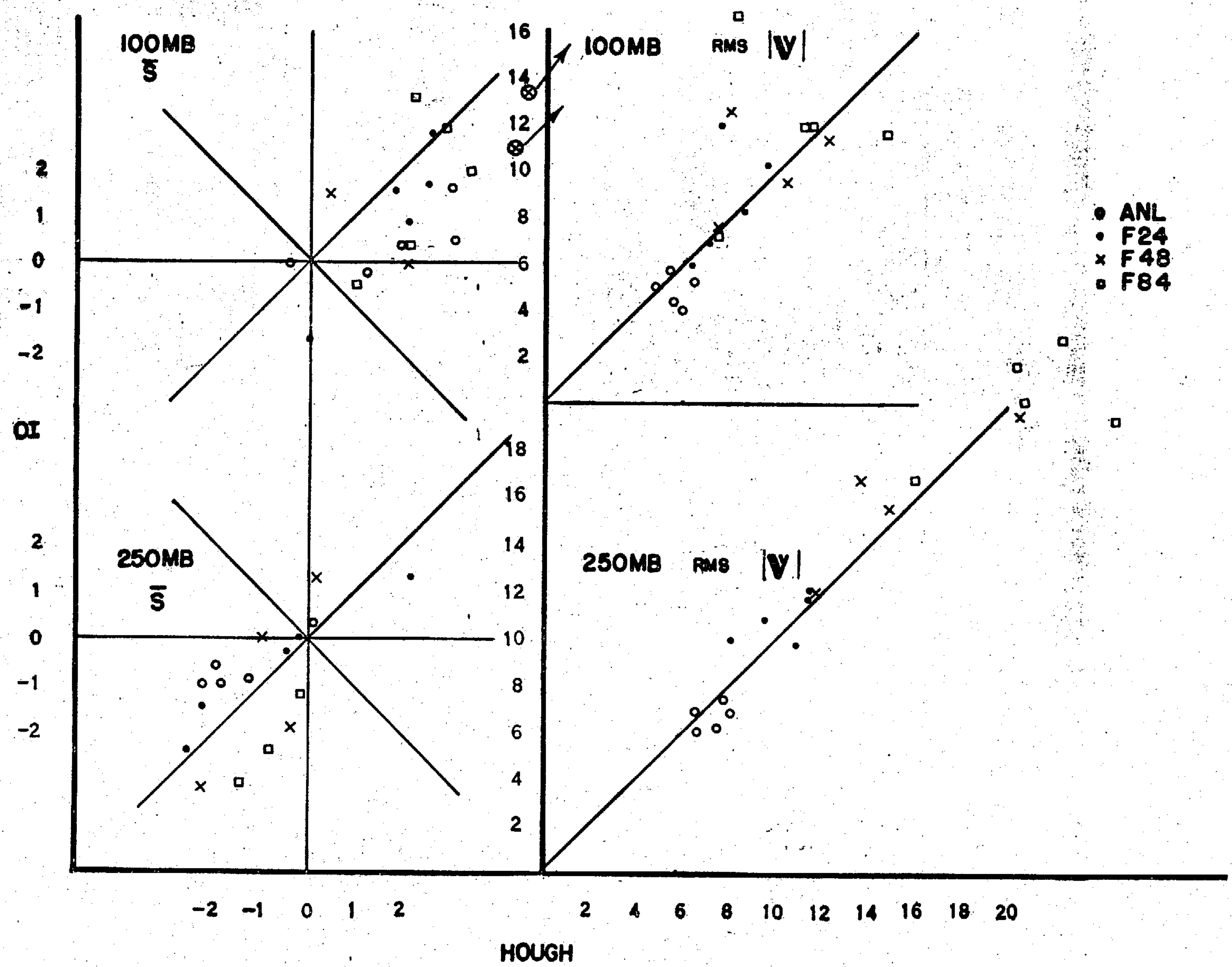

NA 110
TEMPERATURE ERRORS
100MB T
250MB T
850MB T
OI
HOUGH
100MB RMS
250MB RMS
850MB RMS
ANL
F24
F48
F84

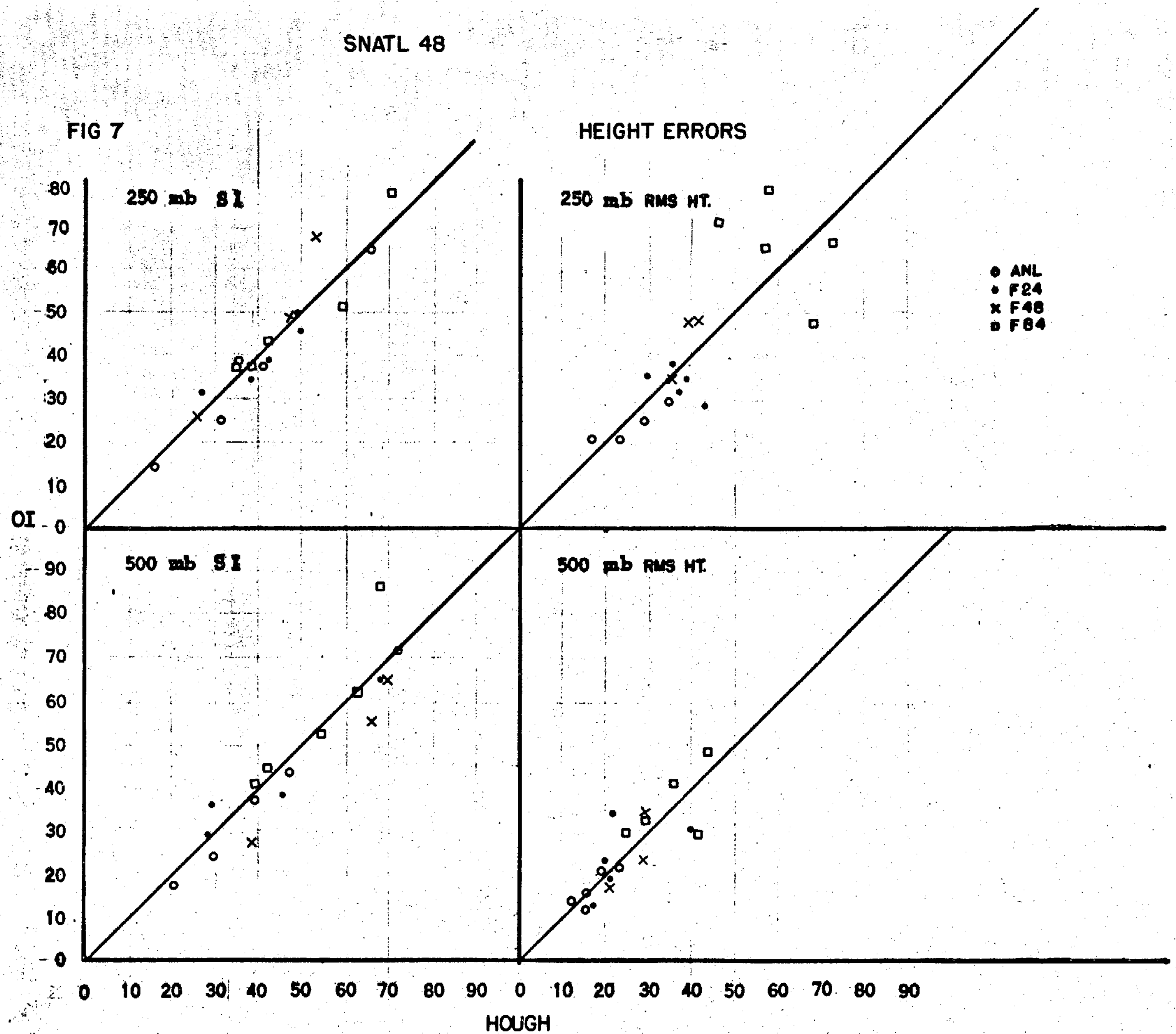
FIG 6


NA110
WIND ERRORS
100MB
$\bar{S}$
100MB
RMS |V|
250MB
$\bar{S}$
250MB
RMS |V|
OI
HOUGH
ANL
F24
F48
F84

SNATL 48

FIG 7

HEIGHT ERRORS



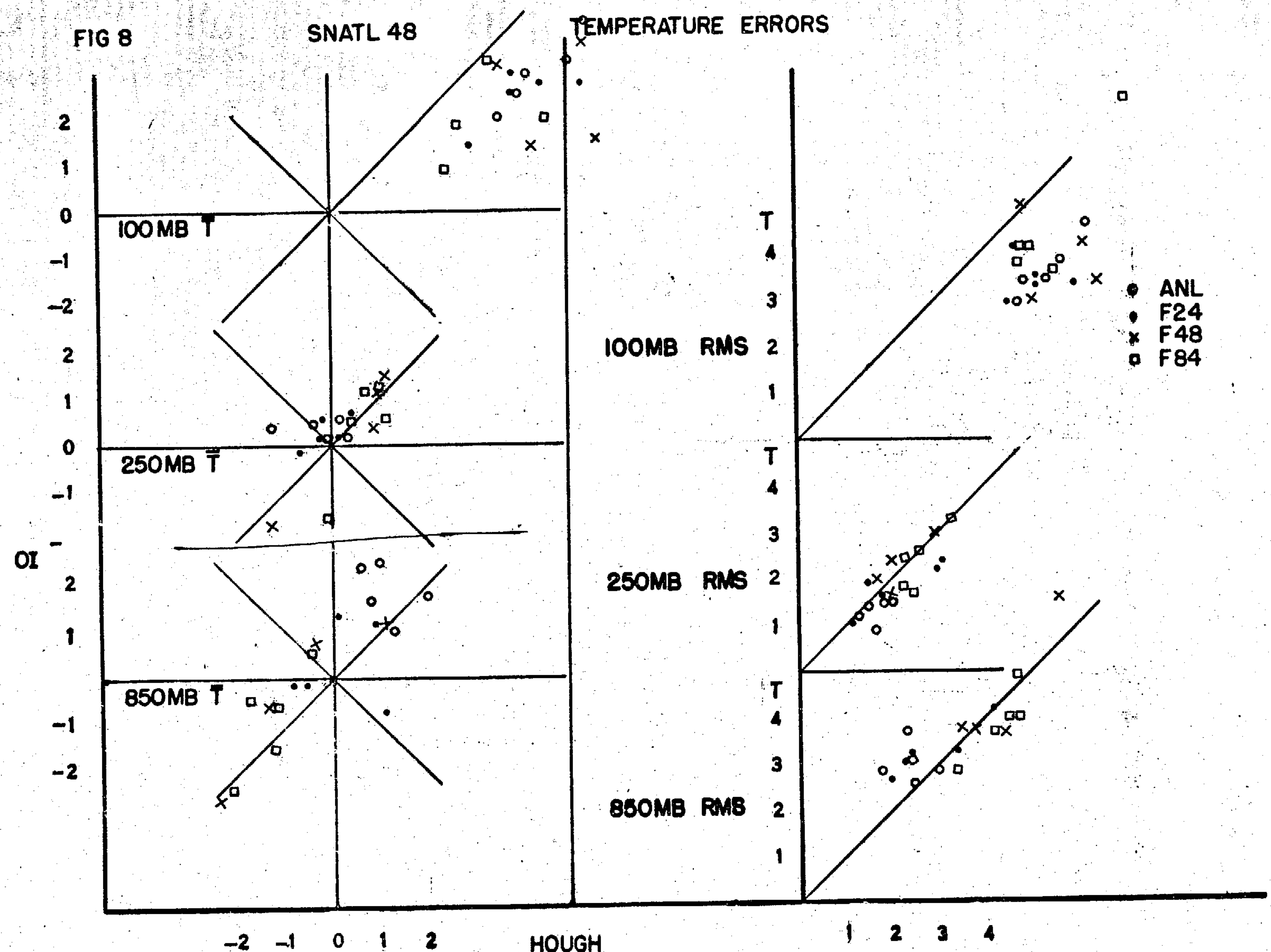
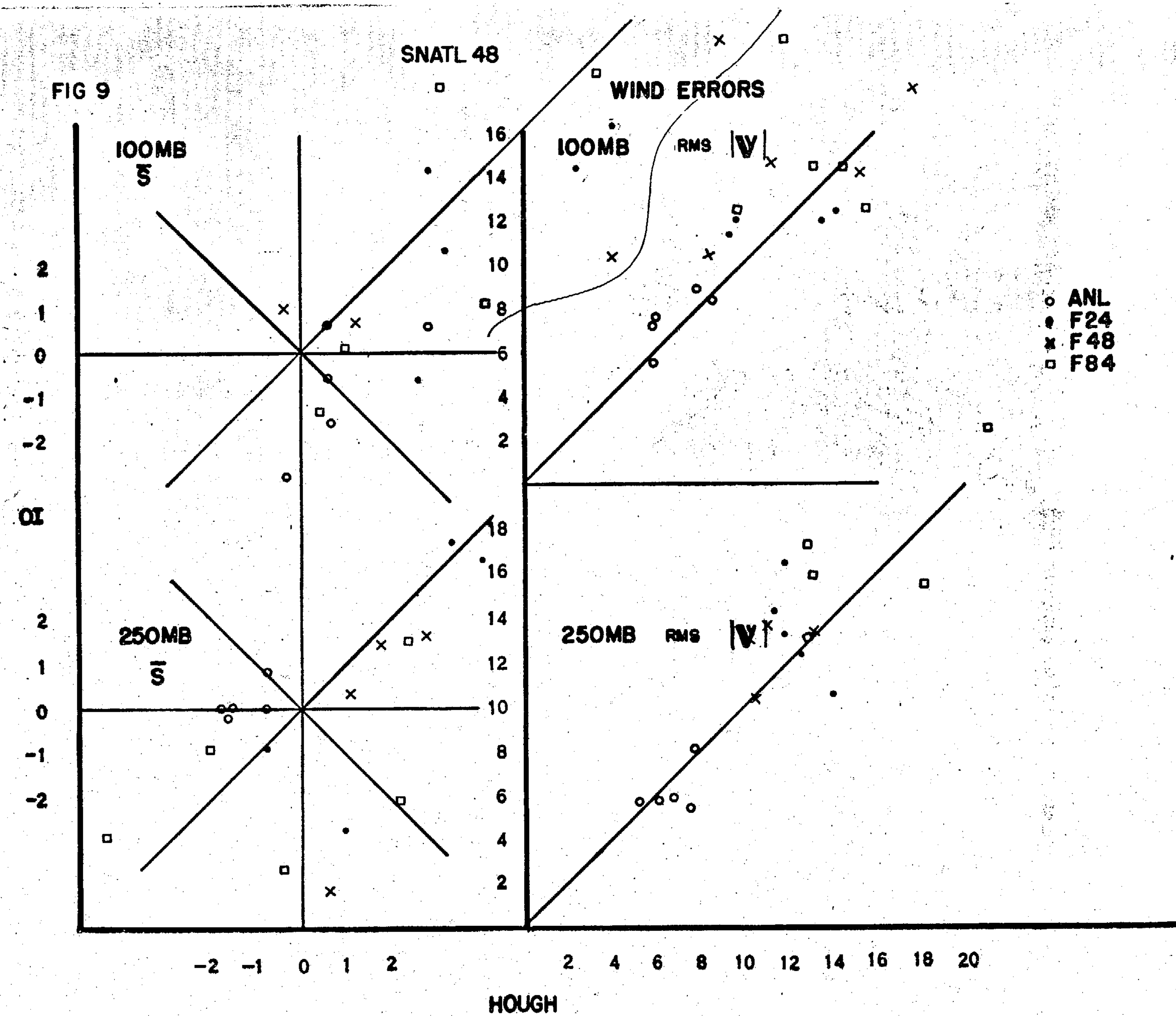


FIG 8
SNATL 48
TEMPERATURE ERRORS
100MB T
250MB T
850MB T
OI
HOUGH
100MB RMS
250MB RMS
850MB RMS
T
ANL
F24
F48
F84
2
1
0
-1
-2
2
1
0
-1
2
1
-1
-2
4
3
2
1
-2
-1
0
1
2
1
2
3
4

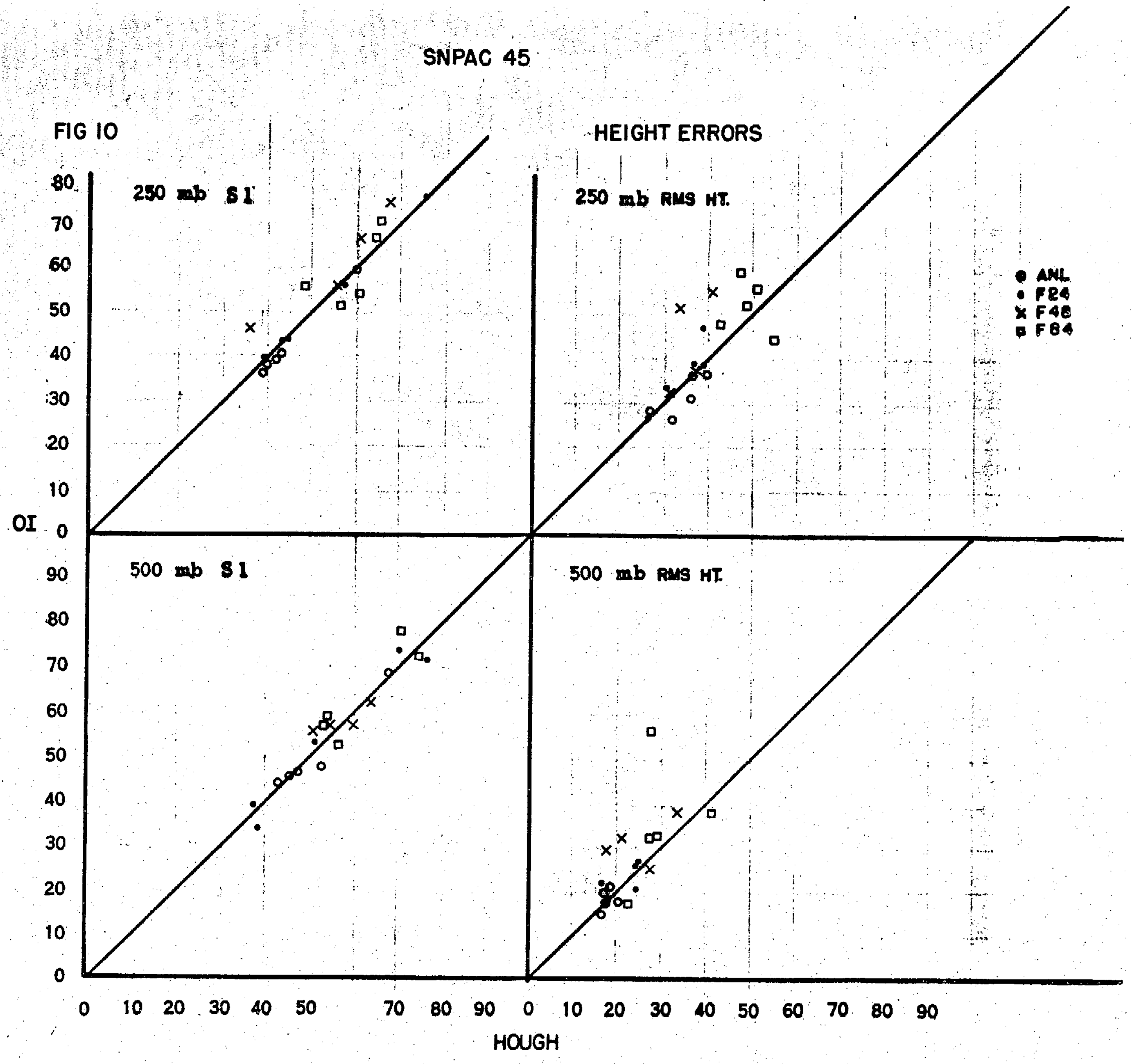


SNATL 48
FIG 9
WIND ERRORS
100MB
$\bar{S}$
100MB
RMS |V|
250MB
$\bar{S}$
250MB
RMS |V|
ANL
F24
F48
F84
OI
HOUGH

SNPAC 45

FIG 10

HEIGHT ERRORS



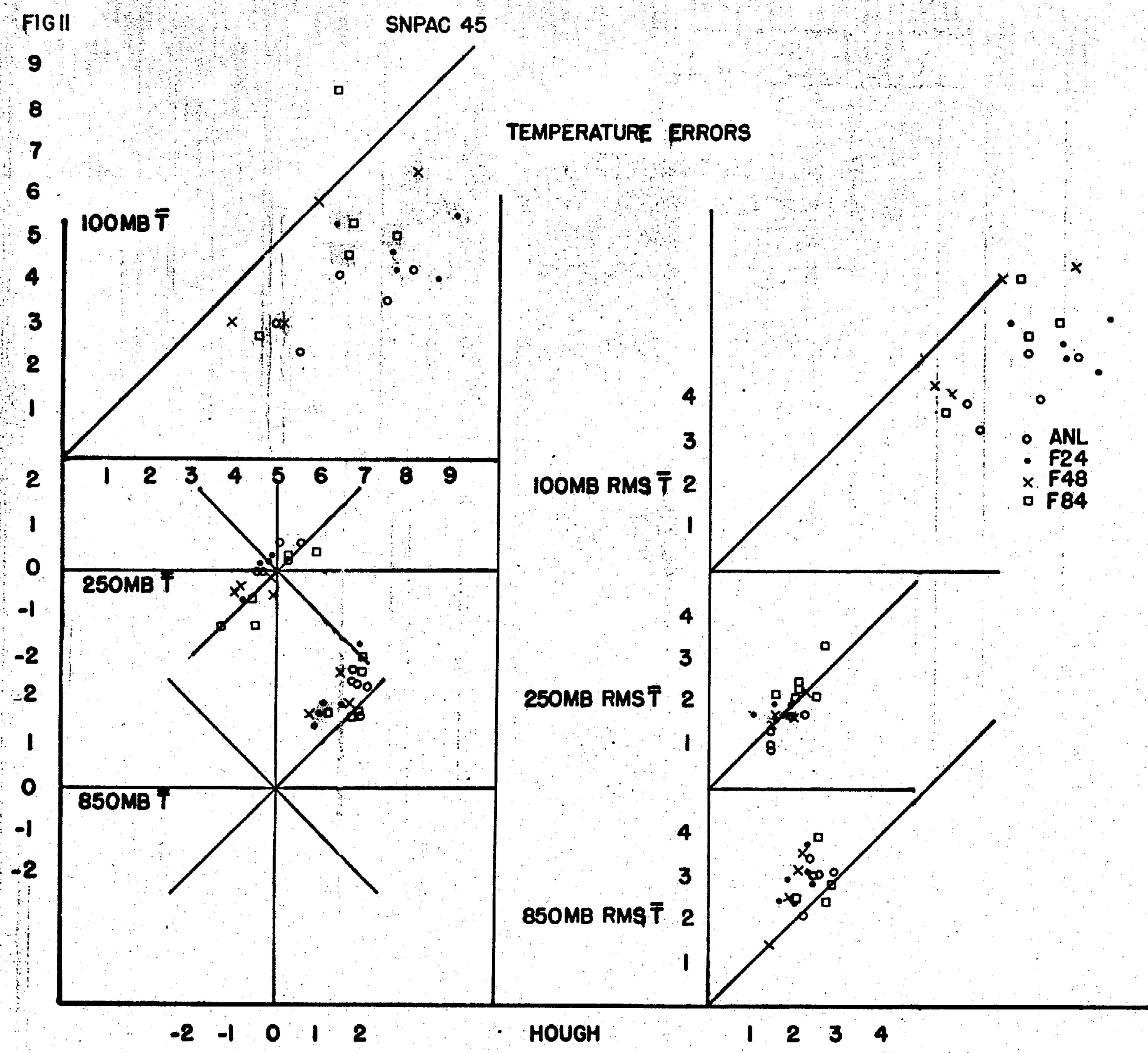
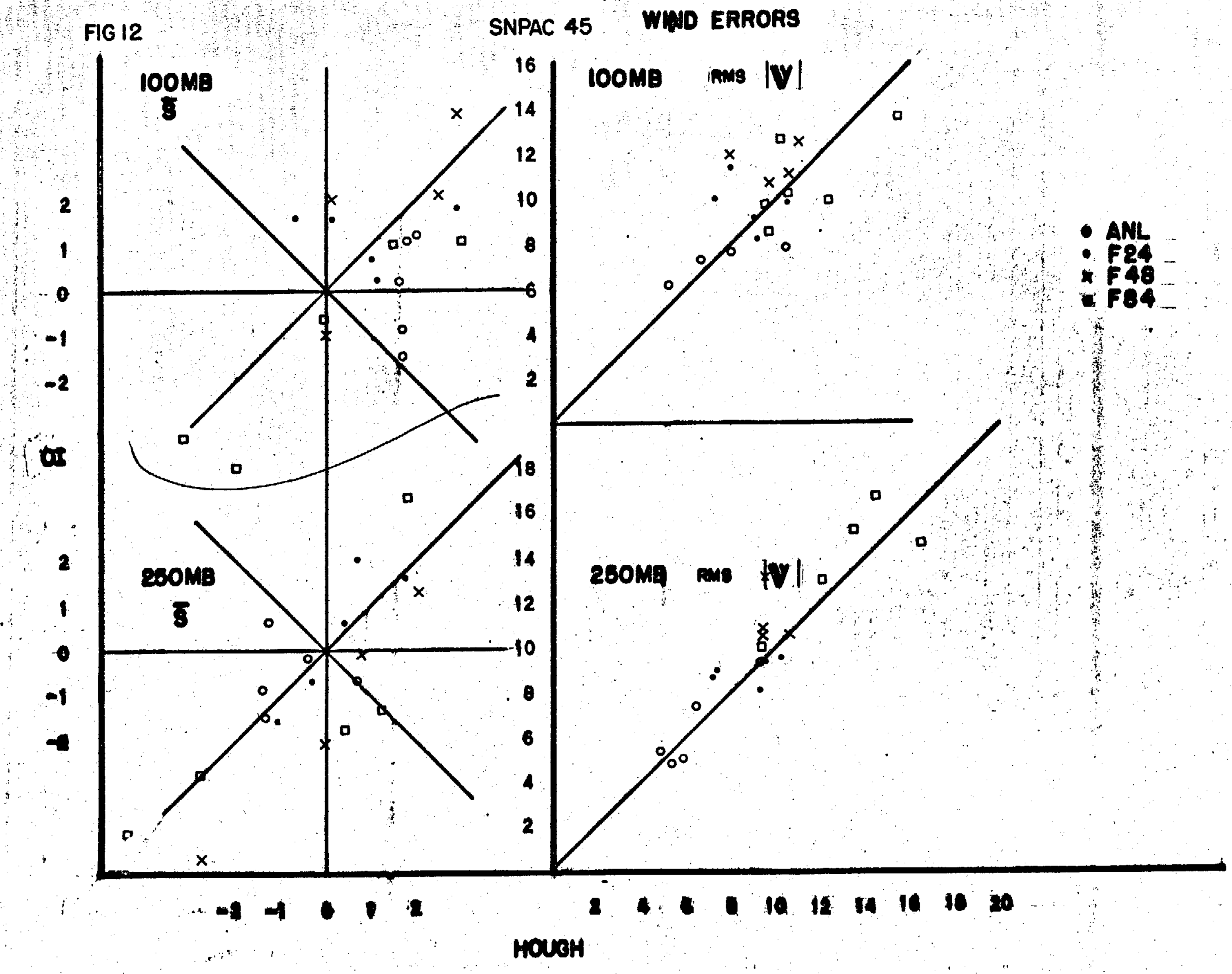


FIG II
SNPAC 45
TEMPERATURE ERRORS
100MB T̄
250MB T̄
850MB T̄
100MB RMS T̄
250MB RMS T̄
850MB RMS T̄
HOUGH
ANL
F24
F48
F84

FIG 12



SNPAC 45
WIND ERRORS
100MB
$\bar{S}$
100MB RMS |V|
250MB
$\bar{S}$
250MB RMS |V|
ANL
F24
F48
F84
HOUGH

FIG 13 SH 31 HEIGHT ERRORS

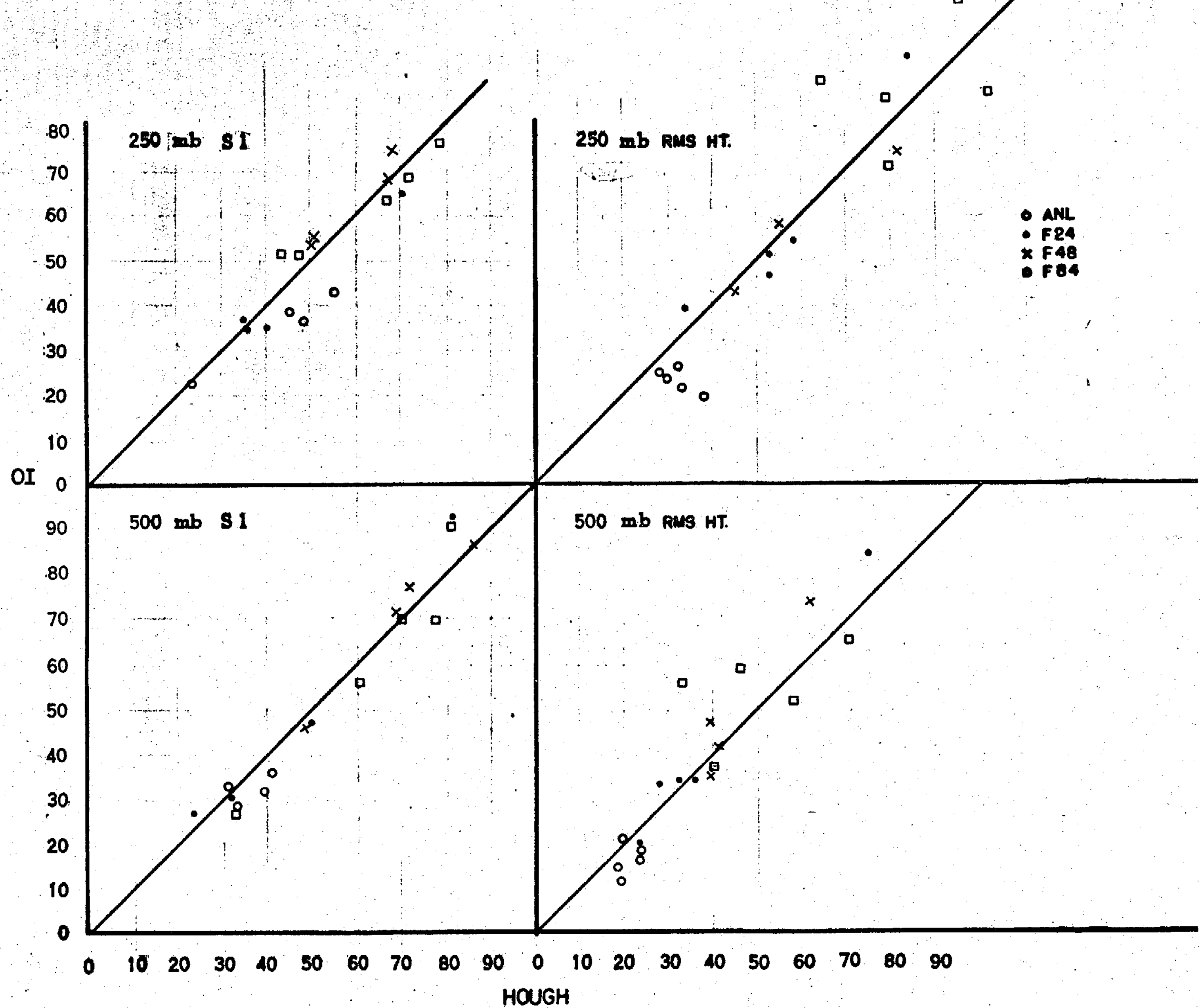
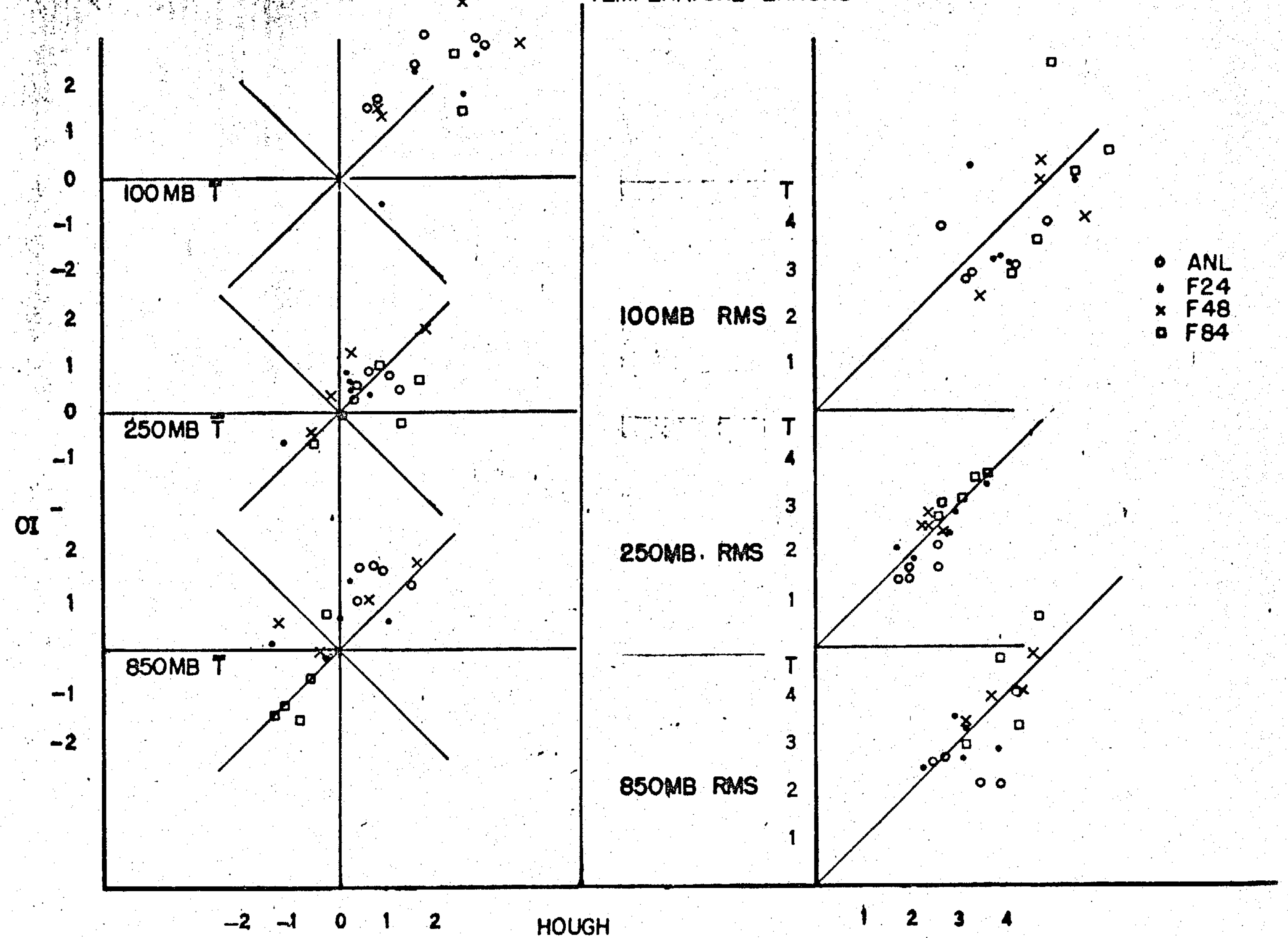


FIG 14 SH 31 TEMPERATURE ERRORS



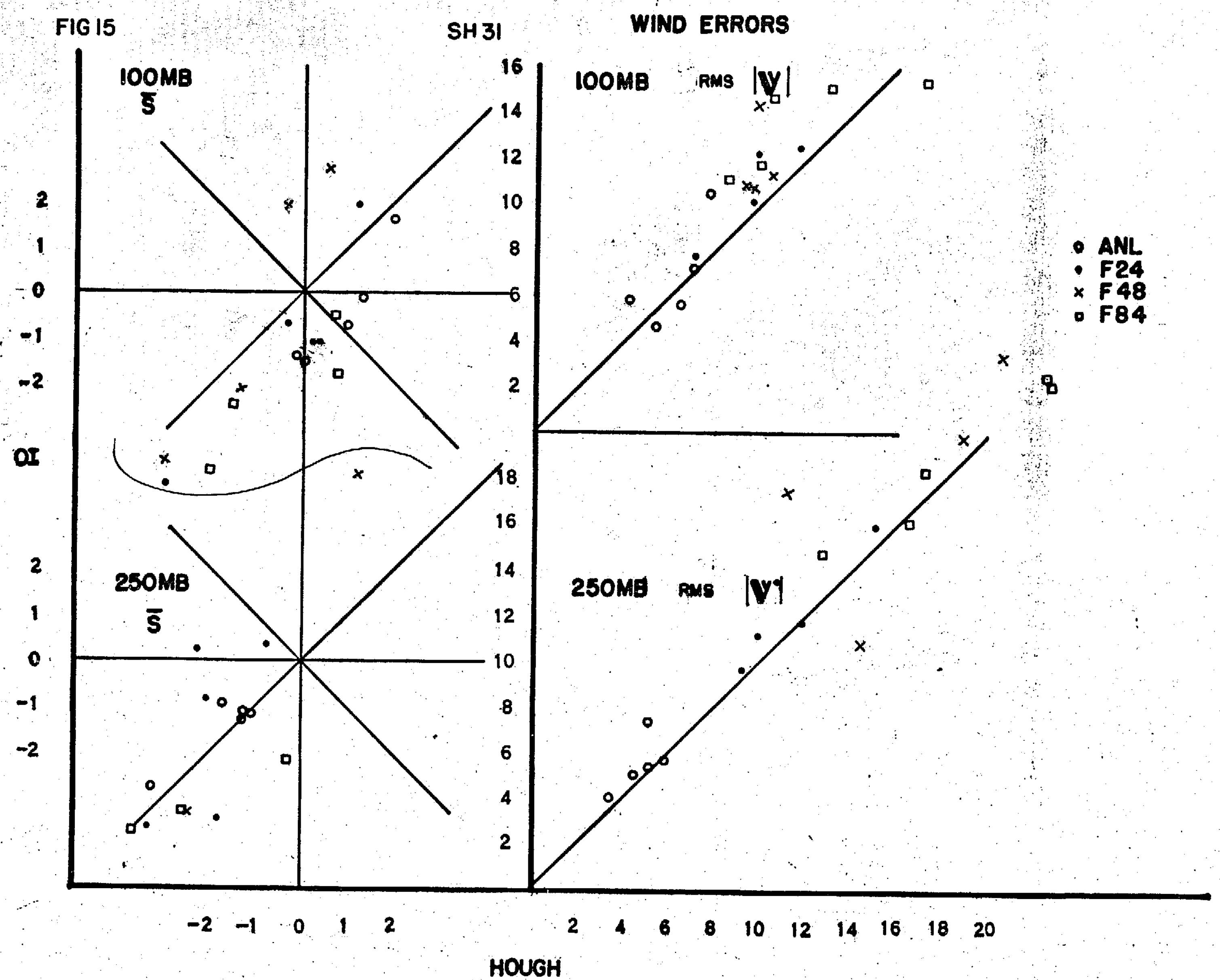
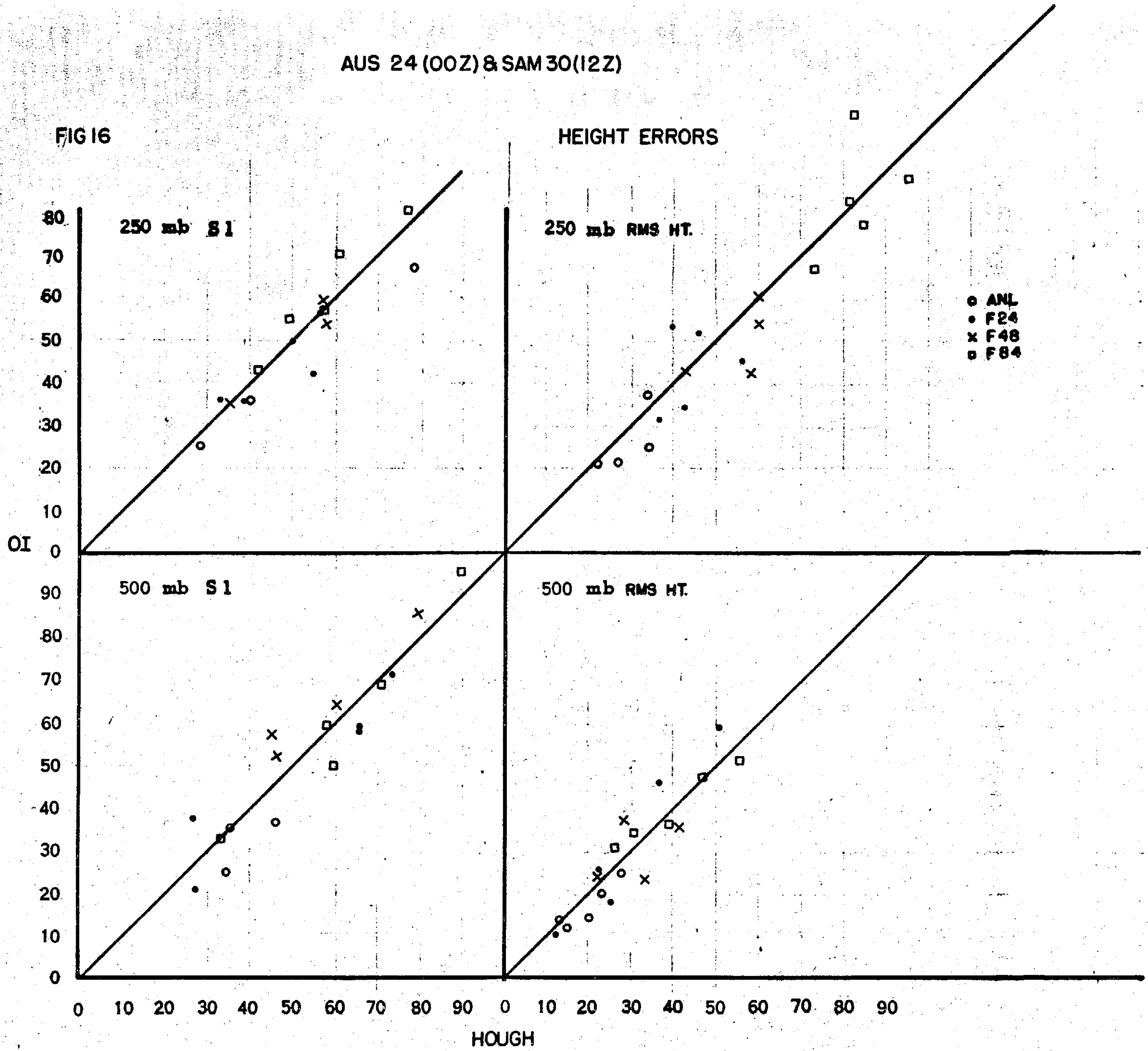


FIG 15
SH 31
WIND ERRORS
100MB
$\bar{S}$
100MB RMS $|V|$
250MB
$\bar{S}$
250MB RMS $|V|$
ANL
F24
F48
F84
OI
HOUGH

AUS 24 (00Z) & SAM 30(12Z)

FIG 16

HEIGHT ERRORS



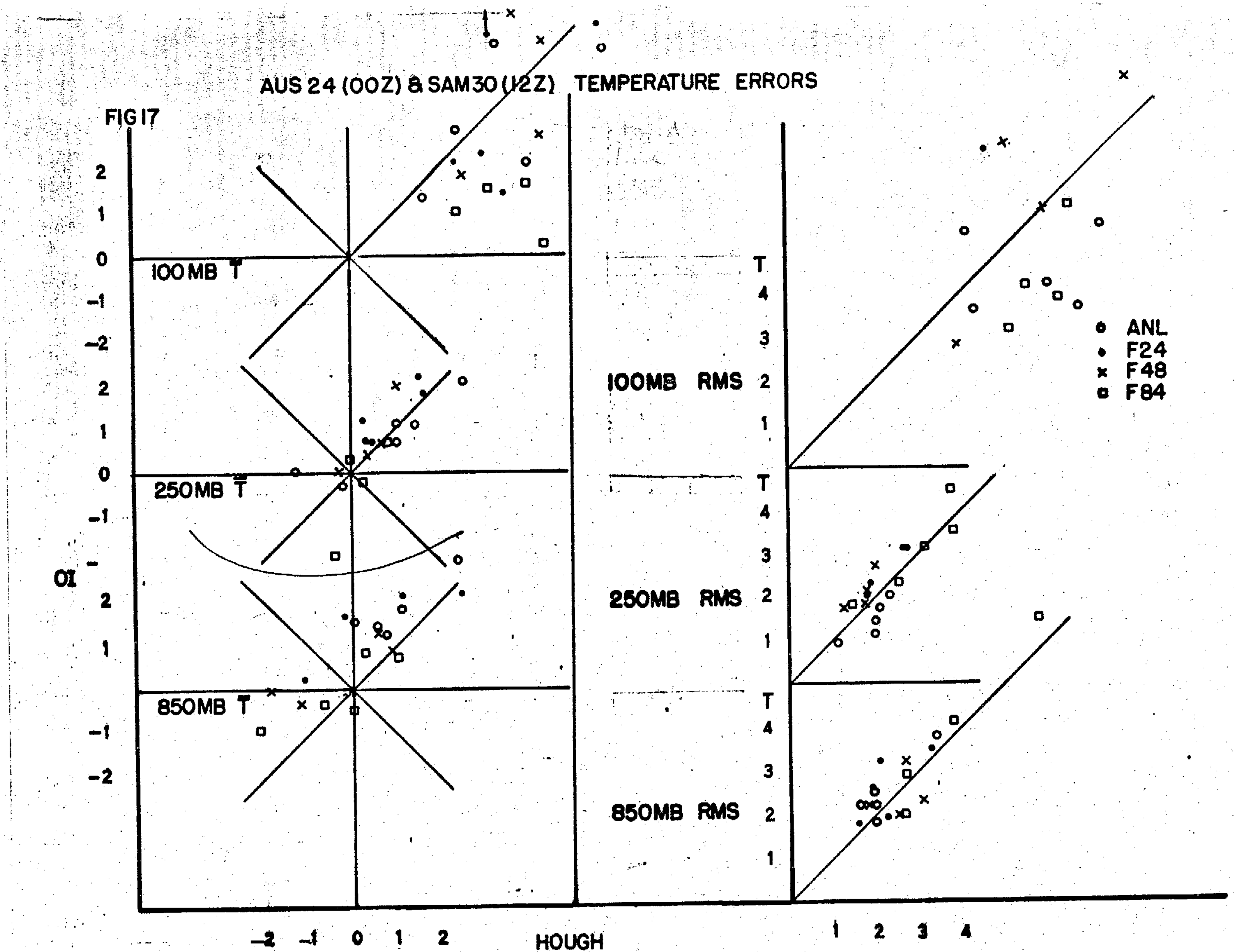
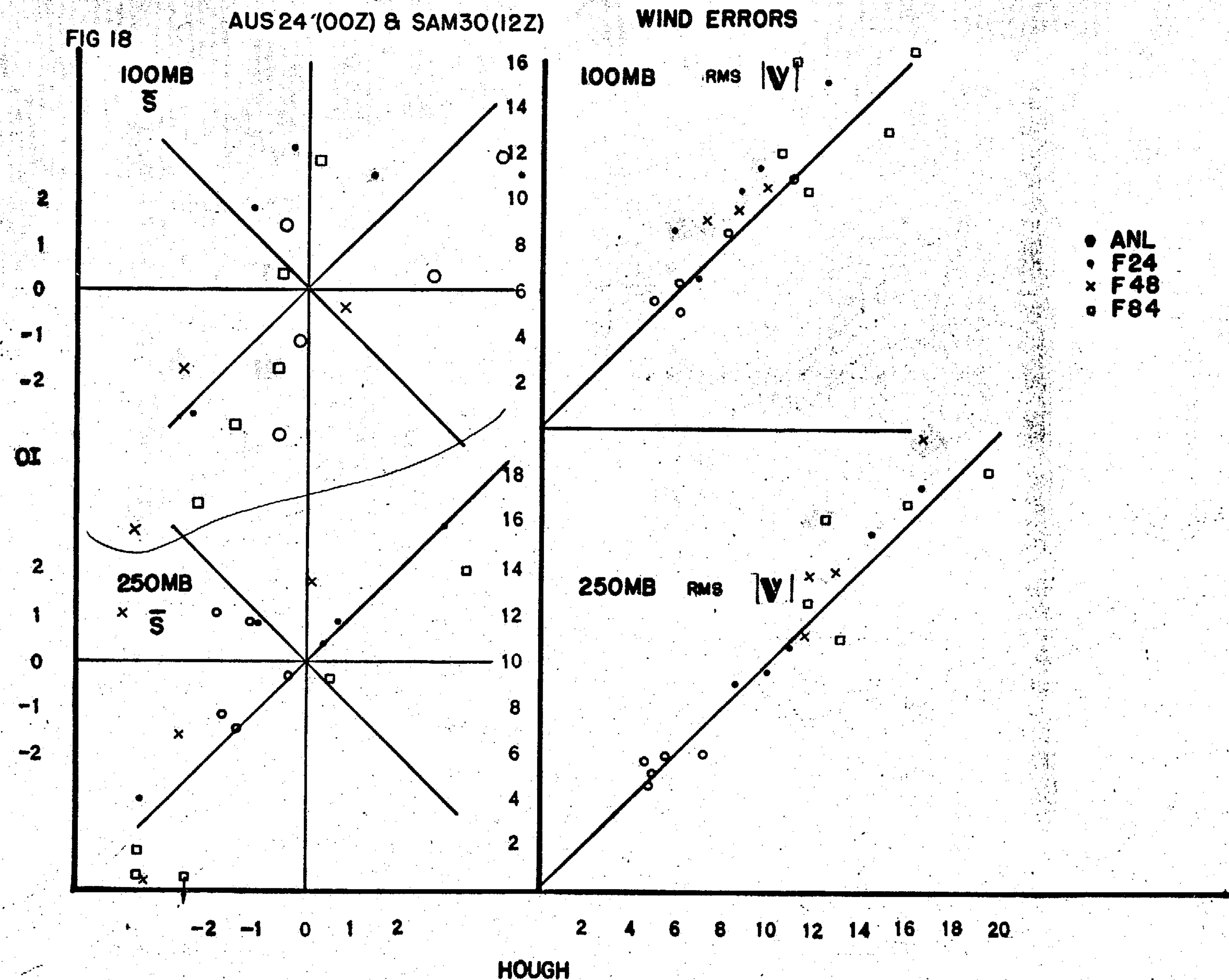

AUS 24 (00Z) & SAM30 (12Z) TEMPERATURE ERRORS
FIG 17
100MB T̄
250MB T̄
850MB T̄
OI
HOUGH
100MB RMS
250MB RMS
850MB RMS
T
ANL
F24
F48
F84

FIG 18



AUS 24 (00Z) & SAM30 (12Z)
WIND ERRORS
100MB
$\bar{S}$
100MB RMS |V|
250MB
$\bar{S}$
250MB RMS |V|
OI
HOUGH
● ANL
• F24
× F48
▫ F84

OI - HUF TEST: SMG3C 12-72HRS (5 Cases), 84HRS (4 Cases); HUF (X), OI (●), OBSVD (Col)

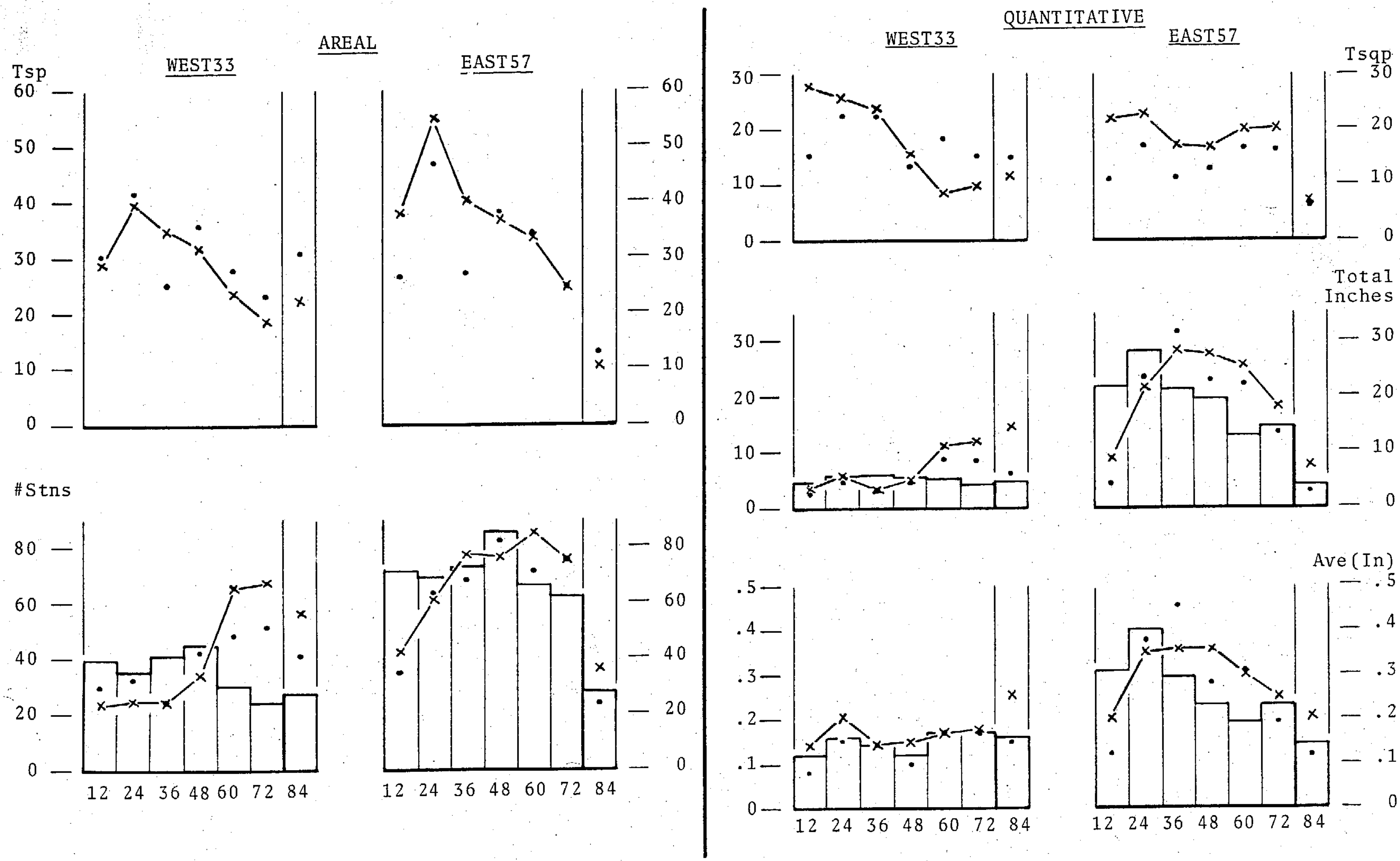
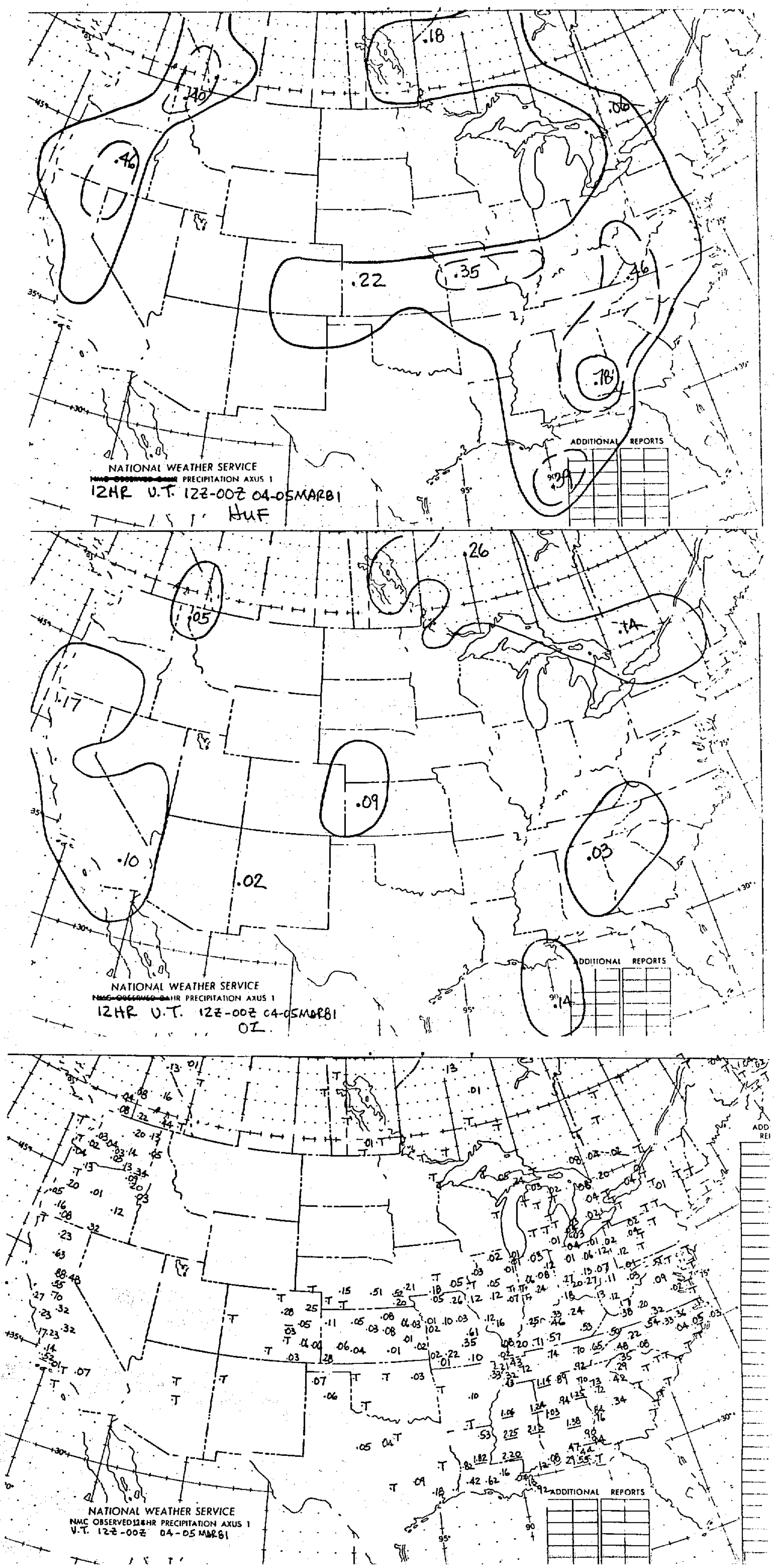
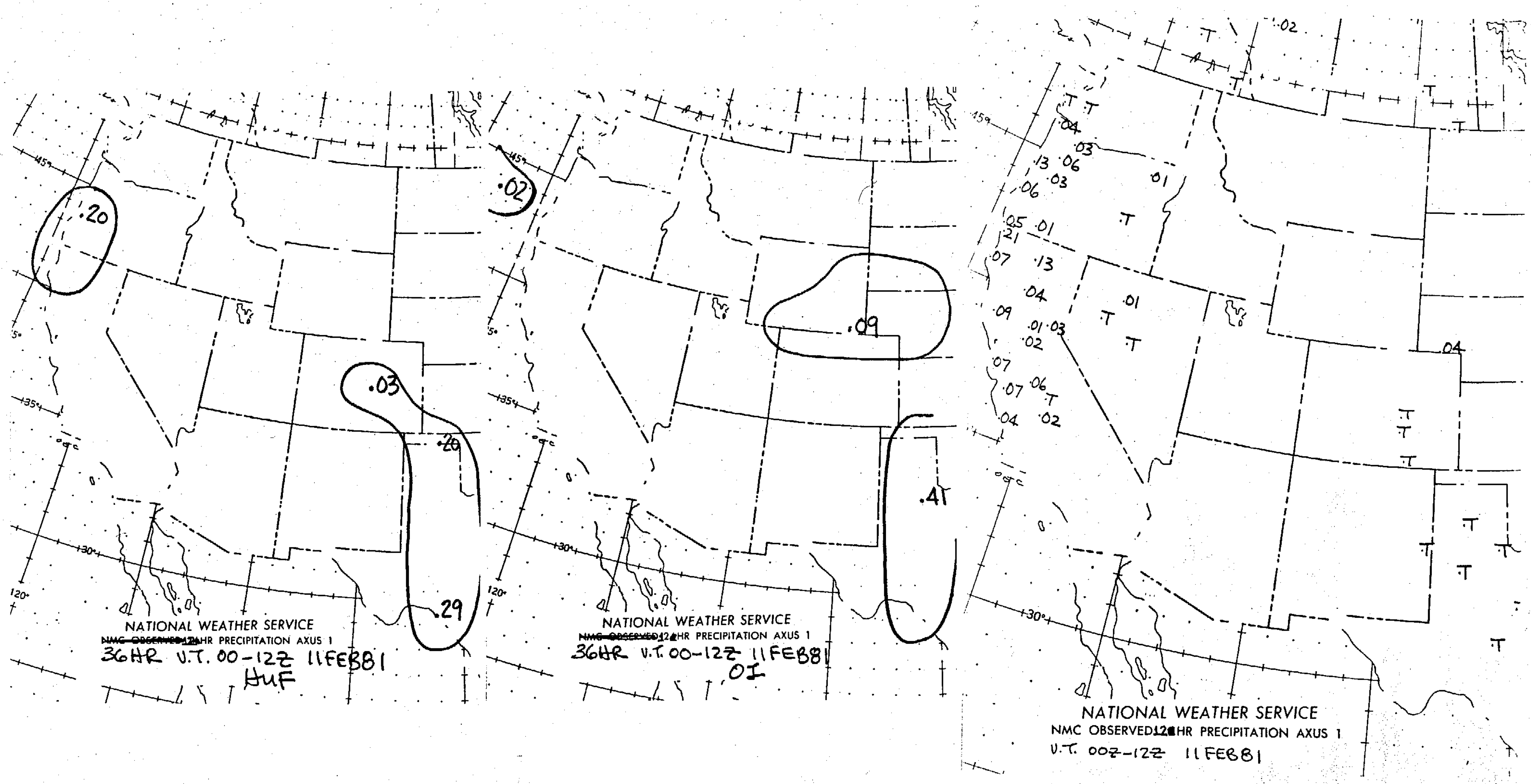


FIG19



.18
.40
.46
.06
.22
.35
.46
.78
.39
ADDITIONAL REPORTS
NATIONAL WEATHER SERVICE
~~NMC OBSERVED 24HR~~ PRECIPITATION AXUS 1
12HR U.T. 12Z-00Z 04-05MAR81
HUF
.26
.05
.14
.17
.09
.10
.02
.03
.14
ADDITIONAL REPORTS
NATIONAL WEATHER SERVICE
~~NMC OBSERVED 24HR~~ PRECIPITATION AXUS 1
12HR U.T. 12Z-00Z 04-05MAR81
OI
ADDITIONAL REPORTS
NATIONAL WEATHER SERVICE
NMC OBSERVED 12HR PRECIPITATION AXUS 1
V.T. 12Z-00Z 04-05MAR81

FIG 20



.20
.03
.20
.29
NATIONAL WEATHER SERVICE
~~NMC OBSERVED~~ 12 HR PRECIPITATION AXUS 1
36HR V.T. 00-12Z 11FEB81
HUF
.02
.09
.41
NATIONAL WEATHER SERVICE
~~NMC OBSERVED~~ 12 HR PRECIPITATION AXUS 1
36HR V.T. 00-12Z 11FEB81
OI
NATIONAL WEATHER SERVICE
NMC OBSERVED 12 HR PRECIPITATION AXUS 1
V.T. 00Z-12Z 11FEB81

FIG 21

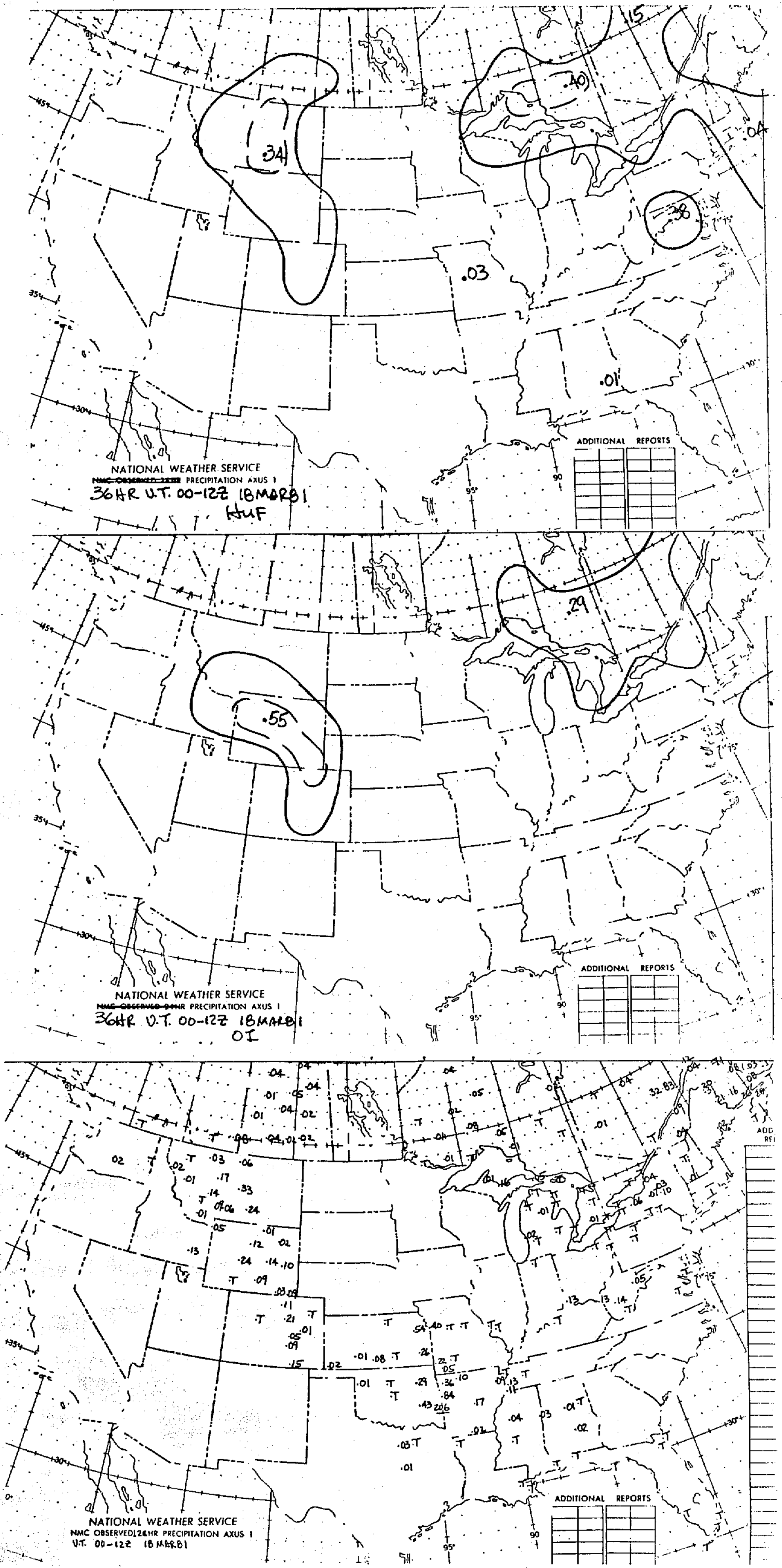

.34
.40
.15
.04
.38
.03
.01
ADDITIONAL REPORTS
NATIONAL WEATHER SERVICE
~~NMC OBSERVED 24HR~~ PRECIPITATION AXUS 1
36HR U.T. 00-12Z 18MAR81
HUF
.29
.55
ADDITIONAL REPORTS
NATIONAL WEATHER SERVICE
~~NMC OBSERVED 24HR~~ PRECIPITATION AXUS 1
36HR U.T. 00-12Z 18MAR81
OI
ADDITIONAL REPORTS
NATIONAL WEATHER SERVICE
NMC OBSERVED 24HR PRECIPITATION AXUS 1
U.T. 00-12Z 18MAR81

FIG 22

Appendix A

SUBJECTIVE EVALUATION

OI vs HOUGH

as initial conditions for

the Operational SMG3C

Instructions

Please indicate by symbol:

HUF

O/I

T (for tied)

which analysis system gave rise to the better analysis or forecast in your field of specialization. Feel free to make comments wherever you can find space to do so. Please don't bother with distinctions such as "slightly" or "much" better - if the "improvement" is borderline mark it a T (tie).

Initial Date/Time 12Z 30 Jan 81

Evaluator ______________________

Specialization Area

Precipitation D. Olson

	24 hr.	48 hr.
Rain/No Rain Coverage		
Western (West of 105°) U. S. -------	☐	☐
Eastern U. S. ---------------------	☐	☐
Quantitative Amounts		
Western U. S. ---------------------	☐	☐
Eastern U. S. ---------------------	☐	☐
Relative Humidity Patterns		
Western U. S. ---------------------	☐	☐
Eastern U. S. ---------------------	☐	☐
Utility of Mass/Motion Forecasts to QPF		
All U. S. ------------------------	☐	☐

Specialization Area

Basic Weather D. Saxton

	24 hr.	48 hr.
SLP & Thickness		
East Pacific & Alaska ------------	☐	☐
Western (West of 105°) U. S. ------	☐	☐
Eastern U. S. --------------------	☐	☐
Western Atlantic -----------------	☐	☐
500 mb Height & Vorticity		
East Pacific & Alaska -------------	☐	☐
Western U. S. ---------------------	☐	☐
Eastern U. S. ---------------------	☐	☐
Western Atlantic ------------------	☐	☐

Specialization Area

Aviation R. McCarter

	Anal	24 hr.
250 mb Hts and Isotachs (and Jet)		
U. S.	☐	☐
Pacific	☐	☐
Tropics (wherever you please)	☐	☐
South America	☐	☐
Australia	☐	☐
100 mb Hts & Isotherms		
Pacific	☐	☐
U. S.	☐	☐
Atlantic	☐	☐
South America	☐	☐
Australia	☐	☐
Tropopause Pressure		
Pacific	☐	☐
U. S.	☐	☐
Atlantic	☐	☐
Tropics (wherever)	☐	☐
South America	☐	☐
Australia	☐	☐
Tropopause Vertical Wind Shear		
Pacific	☐	☐
U. S.	☐	☐
Atlantic	☐	☐
Tropics (wherever)	☐	☐
South America	☐	☐
Australia	☐	☐

Specialization Area

Aviation R. McCarter

Extra Credit:

The { HUF based [ ] / O/I based [ ] / Neither [ ] } (check one)

forecast should replace the current persistence "forecast" in the southern hemisphere (ATA use primarily).